Process Diagnosis System (PDS) – A 30 Year History

E. D. Thompson\textsuperscript{a}, E. Frolich\textsuperscript{a}, J. C. Bellows\textsuperscript{a}, B. E. Bassford\textsuperscript{a}, E.J. Skiko\textsuperscript{a}, M. S. Fox\textsuperscript{b}

\textsuperscript{a}Siemens Energy, Inc. Orlando, Fl. 32826
\textsuperscript{b}Mechanical & Industrial Engineering, University of Toronto

Abstract
PDS (Process Diagnosis System) is an expert system shell developed in the early 1980's. It could handle thousands of sensor inputs and produce thousands of diagnostic messages with confidence factors based on complex logic designed to mimic the thinking of human experts. PDS went into commercial operation in 1985 to monitor seven power plant generators from a centralized diagnostic center at Westinghouse Power Generation headquarters. In the 1990's the popularity of advanced technology gas turbines provided a renaissance in PDS utilization. The software has undergone rewrites and improvements since its inception, and the current PCPDS now supports the Siemens Power Diagnostics\textsuperscript{®} Center with centralized rule based monitoring of over 1200 gas turbines, steam turbines, and generators.

Introduction

In 1981, Mark S. Fox, a Research Scientist and Director of the Intelligent Systems Laboratory of the Robotics Institute at Carnegie Mellon University (CMU) was invited by Westinghouse to give a seminar on Artificial Intelligence. After the seminar, a Westinghouse manager pulled Mark aside. Westinghouse, at that time, manufactured steam turbines and generators (herein referred to as units) that were used in power generation plants around the world. The question that the Westinghouse manager asked was could Expert Systems technology be used to model their engineers' expertise to detect and/or predict the pending occurrence of a malfunction with sufficient advanced warning to allow the scheduling and execution of maintenance to the extent that the malfunction and associated forced outage could be avoided. By early diagnosis of a malfunction a reduction in both the repair costs and downtime can be realized. Thus began the story of PDS, originally known as Process Diagnosis System and today PCPDS, defined as Personal Computer Process Diagnosis System. PDS was launched as a production service in 1985 and is still in use today monitoring over 1,200 turbines and generators around the world.

In this paper we trace the authors’ recollections of the history of the development of PDS, both the expert systems "engine" and knowledge base, and describe some changes in the global marketplace and their impact on PDS.

In the Beginning at Westinghouse

In the late 1970’s, Westinghouse concluded that the next major development in control systems would be in the area of diagnosing potential issues because turbine-generators in commercial power plants could take weeks to repair. Westinghouse produced a diagnostic algorithm with three inputs and three outputs based on conditional probabilities. Since there were no proven values for the probabilities, experts estimated the values. It was clear to Westinghouse that the algorithm would not scale to a larger number of sensors and potential issues characteristic of a steam turbine generator system. The challenge was to see if Expert Systems could solve the turbine-generator predictive diagnosis problem.

Turbines and generators are large, complicated devices that are engineered to operate reliably for extended periods of time. It is not uncommon to have over 1,000 sensors on this equipment measuring temperature, pressure, vibration, etc. to control and protect their operation. This instrumentation was designed to generate readings (e.g. once a second) on a continuous basis. The control system is designed to generate an alarm and/or interrupt equipment operation based on the sensor readings, for example, exceeding a threshold. However, the control system had to do this without unduly hampering the operator so subtle changes in sensor readings were not its focus. The question that had to be answered was whether an analysis of sensor readings could detect subtle deviations in the sensor readings (prior to exceeding typical engineering designated thresholds) that could in turn be used to predict failures. Mark Fox interviewed Westinghouse power systems experts to understand how they would diagnose power plant equipment based on the available sensor indications. It was eventually determined that equipment failures could be predicted and diagnosed based on subtle deviations in sensor readings.

The knowledge engineering process revealed a number of interesting characteristics to the expert systems community of that time:

Used with permission of Siemens Energy, Inc.
1. The diagnosis system could be receiving over 1,000 sensor readings per second per unit being monitored.
2. The sensors fail more often than the unit itself. They either degrade over time or randomly provide spurious data due to exogenous factors. Hence, in addition to diagnosing the units, an effective system would need to also diagnose all of the sensors being used to diagnose the unit.
3. A unit contains many redundant sensors. Hence unit diagnosis may proceed even when it is known that a subset of sensors has failed. The diagnosis has to take the sensor failures into account when deriving its conclusions.
4. The knowledge used to diagnose a unit during "start-up", a transient condition, is very different than during its operational phase, typically a steady-state condition.

Based on the interviews, Fox derived a set of requirements:

1. The goal of the system is to detect symptoms and make a diagnosis before equipment operation is impacted. Hence the system should detect symptoms which indicate a malfunction will eventually occur, as opposed to diagnosing the cause after the malfunction occurs.
2. The expert system cannot operate by looking at the output of a single reading of the sensors, but must base its diagnosis on a series of sensor readings in order to detect subtle changes that occur over time.
3. In order to detect sensor failures, it will need to incorporate sensor analysis techniques, including averaging, voting, and time series analysis.
4. As sensors fail, the diagnosis process should continue to diagnose the unit with whatever information continues to be transmitted.
5. The expert system will have to be context aware and use different knowledge based on the operating state of the unit.

In 1981, there were two types of expert systems. The first utilized the approach of backward chaining rules with a method for propagating belief / certainty (aka Certainty Factors). MYCIN (Shortliffe, 1974), a medical diagnosis system, was the best known example. The second utilized the approach of forward chaining rules, called Production Systems (Newell & Simon, 1972), with a number of examples documented in (Rychener, 1976). Detecting unit failures is essentially a diagnosis task for which the MYCIN rule architecture and Certainty Factors were well suited. The difference from typical backward chaining expert systems being that PDS was not diagnosing the cause of a failure after it occurs, but attempting to predict a failure before it occurs. Given the arrival of sensor data every second, Fox combined the forward chaining capabilities of production systems, with the rule architecture and Certainty Factors of MYCIN. The rule base was compiled into a network where sensor readings entered one end as sensor nodes, and certainty measures were propagated along arcs which represented rules to intermediate nodes (i.e., hypotheses) until all of the failure mode nodes were assigned a Certainty Factor. Backward chaining and explanations were also implemented.

In order to address the requirements, a number of unique features were introduced (Fox et al., 1983). With each unit having over 1,000 sensors, rules were not constructed individually to diagnose each sensor. Instead, the knowledge representation language SRL (Wright & Fox, 1982), used to implement the system, represented classes of sensors and rules which were inherited by the instances. A library of rules could be created for a type of unit and instantiated with little effort.

All rules have associated with them a context that specifies under what operating condition the rule is to be used. For example, if the unit is being started, only rules whose context is "start" would be active and used to predict a future failure.

Traditionally, sensor data is processed external to a monitoring system with any suspect data either being modified or removed before the system sees it. Time series analysis, such as data including rate of change, averaging, filtering and smoothing, is used. It was felt that preprocessing sensor data excluded information that may be valuable during the diagnosis process. For example, if it was believed that a sensor is degrading but the readings it generated may still be useful, it might be desired to alter its role in other rules. Consequently, it was decided to process all sensor data within the rule base. To support time series analysis, PDS was designed to provide the capability to store and analyze successive readings of sensors, or values of any other nodes in the rule network using "storage" nodes. Storage nodes stored data over time and specified a function to analyze the data.

Another type of rule, introduced with storage nodes, was a "reading-transform." A reading-transform linked sensor nodes to storage nodes, providing the capability to transform the data that passes through it. Reading-transform rules were used to convert engineering units, scale numbers, etc.

In addition to time series analysis provided by a storage node, PDS introduced a "composite-sensor" node that combined the readings of multiple, overlapping sensors. The node specified a set of sensors and/or reading-sets from which it was to read data, and combined it into a reading for a single virtual sensor. Techniques such as voting and auctioneering were applied to the data.

The final requirement was the handling by PDS of failed sensors while maintaining some level of prediction process. One approach to dealing with failing sensors in PDS is to set the Confidence Factor (CF), similar to the certainty factor in MYCIN, of a sensor to zero when it is determined to be failing. If the sensor is part of the left-hand side (LHS) of a rule, setting its CF to zero would have an unintended effect. If the “fuzzy” AND (i.e., taking the minimum CF) is used to combine the evidence in the LHS
of the rule, then the rule's LHS CF would be zero, thereby eliminating any contribution by other evidence in the LHS. If a weighted combination of the evidence in the LHS of the rule was used it too would reduce the impact of the remaining evidence in the LHS of the rule because the weight of the CF of the failing sensor would remain the same. The approach was to introduce the "parametric-alteration" rule. Once sensor diagnosis is complete, the parametric alteration rules are executed. If sensors are found to fail, or are degrading slowly, the parametric alteration rule would modify the weight of the failing sensors in the LHS of rules used to diagnose the unit.

The above briefly describes the architecture of PDS. More details can be found in Fox et al. (1983). It is this architecture that formed the core of the PDS product; a product designed to be robust enough to continue to produce results even with marginal sensor data quality.

The LISP version of PDS that Mark Fox wrote was demonstrated to Westinghouse management on March 20, 1982. James Bellows and others were present at the meeting. Bellows instantly recognized the potential of the system. They decided that, since Westinghouse's Steam Chemistry Engineering section had considerable field data against which an expert system could be tested, a power cycle chemistry diagnostic system would be a good first project. The result was seen as encouraging and the project was continued.

**From Prototype to Product (1983-1985)**

In late 1983, after Westinghouse moved the Steam-Turbine Generator headquarters to Orlando Florida from Pennsylvania, Craig Weeks (current Siemens CEO of the Power and Gas business unit) built the first diagnostic center and gave the artificial intelligence product a visible nature. Robert Osborne, manager of Diagnostics, Monitor and AI Development at Westinghouse Electric Corporation, describes this diagnostics center in detail in his paper titled “On-line Artificial Intelligence-Based Turbine Generator Diagnostics” from 1986. Some 25 plus years before the age of “Big Data” and the “Internet of Things” the tremendous value in acquiring sensor data from operating equipment for the purpose of analysis and diagnostics was recognized.

Also in 1983, Bellows started writing a chemistry diagnostic system for a fossil fired power plant. This product would eventually be called ChemAID (Chemistry Artificial Intelligence Diagnostics). As recognized during the early development of PDS, one of the major issues was verifying the reliability of the plant instrumentation. Instrumentation had to be diagnosed before conclusions about the plant could be drawn. Storage nodes, reading-transforms, composite sensors, parametric alteration rules and Confidence Factors had to be fully utilized to make effective diagnoses.

Over several years, Bellows wrote multiple fossil diagnostic systems, improving each of the previous versions based on new experiences. The penultimate system written by Bellows led to his first paper on a diagnostic system (Bellows, 1984). In the process he noticed that many rules were duplicated, so PDS was given subsystems that could be instantiated multiple times with different suffixes on all the nodes and rules. The trick was to then have nodes in the main system that had the same names as the nodes from the instantiated subsystems. PDS would automatically recognize two nodes with the same name and merge them. This technique was used to instantiate the mathematics package as many as forty times. It was also used to allow the modular construction of the rule base. It was recognized that different boiler designs would need different rule bases. A once-through boiler subsystem rule base was written. A drum boiler subsystem rule-base was also written. The interfaces with other subsystems were designed to make them interchangeable at the PDS level. Initially when a mathematical function or other special operation was needed, the rule writers had to request it from the PDS programmers. Eventually the User Designed Operation (UDO) feature was created so that knowledge engineers could create their own operators. In general, the chemistry development drove changes in PDS so that when a generator diagnostic rule base was written, all the necessary operators were already in PDS. That was not true for TurbinAID (Steam Turbine Artificial Intelligence Diagnostics – a subsequent product) where some special functions were needed that had not been necessary in the other diagnostic systems.

Two important style issues were developed in writing ChemAID. First, Bellows, an expert chemist, wrote the rules so that only one thought operation occurred in a single rule. The diagnostic problem was written as if it were being thought through for the first time. This style made modification significantly more straightforward because only one thing was modified at a time. The second style point was a convention to name nodes descriptively and then to name the rules supporting them as the node name plus a rule number, as in boiler-sodium-high-r1, boiler-sodium-high-r2, etc. This naming convention made the rule base relatively readable and easier to modify if testing showed weaknesses in various diagnoses. When ChemAID was applied to a commercial unit, substantial modification of the rules was required because the sensors were not as precise as the rule base had originally assumed. In effect, the rule base had to be “more flexible”.

Although the chemistry diagnostic systems were the first to be written, the first commercial use of a PDS rule base was on generators. A client with several large fossil plants running at maximum capacity had a goal of maintaining high reliability. Therefore, they wanted to have all the information they could to help mitigate the risk of generator outages. GenAID (Generator Artificial Intelligence Diagnostics), designed to diagnose many potential generator issues using PDS, was installed.

In March 1985, the Westinghouse Diagnostic Operation Center (DOC) went live with real time diagnostics driven by PDS and 24/7 staffing by engineers. Real time in this case was about 98% of the diagnoses returning to the pow-
er plant displays, from the DOC, within 2 minutes of the data transmission from the power plant. Data transmissions, from the power plant to the DOC, occurred within a minute of power plant alarms, or about every 15 minutes for routine data. Seven generators from one client were monitored, with the main goal of identifying cracked stator strands.

The 24/7 engineers were also tasked with developing their own section of the rule base, and maintaining customer satisfaction with their assigned sites. There was daily contact with plant operators, and a considerable comfort level was established so that either party could call the other with questions or to discuss an issue. This seemed to be especially true at night when plant operators knew that there was someone else with generator knowledge that was already awake and available as a resource. VAXmail provided text communications before the age of email. In many cases, a close relationship developed between Westinghouse and plant personnel.

Rule base development was originally done on the LISP version of PDS, which had some challenges. These challenges led to the decision to rewrite PDS in VAX C, which was expected to dramatically increase processing speed.

In 1986 as water cooled generator stators became more prevalent, new maintenance recommendations applied and was expected to dramatically increase processing speed. The generator rule base was ready as the new stator temperature monitoring scheme was introduced. The generator rule base was ready as the new stators were commissioned with a patented stator temperature monitoring algorithm that was expected to be an improvement on the standard alarming scheme of a high/low delta temperature. Control systems at that time did not have the capability to implement this rule based algorithm, so it was only available to Diagnostic Center customers.


After the expansion of the group and the start of commercial operation, the rule bases grew in scope and complexity with the experience gained from running the DOC and feedback from plant operators and generator engineering. The working relationship between the DOC crew and plant operators continued. The link to the OEM (original equipment manufacturer) engineering knowledge was a substantial benefit.

Not all the benefits of the diagnostic service were immediately apparent. After a few years the management at one site travelled to the DOC in Orlando to evaluate the cost/benefit of the diagnostic service. Unlike some other sites, there had been no major generator issues diagnosed, and the justification for the system at this site was therefore not fully apparent. The DOC staff ultimately came up with an analysis of the windage loss curve for the generators. It was calculated that a 1% increase in hydrogen purity increased generator efficiency by an amount that justified the expenditure for the service. There were dozens of diagnoses regarding regulator settings and flow adjustments, and the results of the analyses supported at least a 1% improvement in hydrogen purity as a result of the service performed by the DOC. Just that one potential benefit of the Diagnostic Center apparently convinced the clients that the service was cost effective, as they ordered the service.

Toward the end of the 1980’s the first units with generator (GenAID), steam turbine (TurbinAID), and steam chemistry (ChemAID) rule bases were added to the DOC. All three systems combined for a rule base of over 10,000 rules and hundreds of sensors. The Diagnostic Operation Center and the 24 hour diagnostic services were showcased for almost every utility executive that visited the Orlando site, even if it was just the “10 minute tour”. Hundreds of such tours occurred, and the utility executives seemed genuinely impressed with the technology and accomplishments they saw.

The Lean Years (1990-1999)

At the beginning of the 1990’s technical success with providing remote generator diagnostics, combined with improved engineering solutions for common power plant issues of the 1980’s appeared to have increased the reliability of the DOC monitored units.

With the deregulation of the electric utility market, many power companies postponed major service and cut costs to deal with the uncertainty of future business developments. The Westinghouse Power Generation business reacted accordingly.

DOC contracts were phased out in the mid-1990’s. With that came a customer request to purchase standalone systems that would reside at the site and have no yearly service fee. Agreement was reached and a PC based version of PDS was written in C to run on powerful Pentium I computers at the power plant site. PCPDS (Personal Computer Process Diagnosis System) was born. GenAID was subsequently implemented on the PCPDS platform. Because of the way PCPDS was coded, it was even more flexible than the previous PDS. Practically any statement could be put inside rule fields, making self-modifying meta-rules easier to create.

In 1998, Siemens Westinghouse Power Corporation acquired this diagnostic business, along with other assets of the former Westinghouse Power Generation Business. The Diagnostics group delivered a total of 13 standalone systems in the late 1990’s and early 2000’s. During this time, the group was consulted regarding monitoring of the growing gas turbine fleet, with the goal to continue to improve reliability and cost effectiveness of those gas turbines. An attempt to replace PDS with a commercially available expert system shell that would not require software engineering support was made, but nothing was found that was considered capable of handling the complexity of the existing rule base. This resulted in the continued evolution of the PDS, and a period of continual growth until today.
The Growth Years (2000-2013)

The late 90’s were to see tremendous increases in the use of gas turbines for power production. According to the US Energy Information Administration, more than 100 new gas turbines were brought online between 1990 and 1998.

Internal blade cooling technology developed for aircraft turbines in the 1960s had made its way into stationary gas turbines by the late 1970s improving efficiency. In the late 1980s the development of combined-cycle plants yielded further efficiency gains and some specific emissions improvements. In 1985, the expiration of the US Natural Gas Policy Act allowed natural gas to be used for power generation, while rapid discoveries of large natural gas reserves dropped the price of natural gas to just over $2 per million BTU. The Energy Policy Act of 1992 further deregulated the power market, enabling wholesale electrical competition and the rise of energy trading companies. The new, highly competitive energy market could no longer afford the decades-long lead time for constructing new coal plants, nor could it lose time waiting for the typically long power plant start-up times of steam turbines in the face of electrical power outages or blackouts. Drivers such as the US Clean Air Act further led towards the sale of gas turbines for power generation over steam plants.

Unlike the 900°C operating temperatures in steam turbines, the 1400°C turbine inlet temperatures in the advanced technology gas turbines are hot enough to melt parts unless design advancements (e.g. special coatings) were implemented. The more abusive operating conditions that were introduced by the advanced technology gas turbine designs introduced new challenges.

Siemens Westinghouse introduced Long Term Programs (LTP) for their gas turbine customers. These LTPs were intended to provide customers with numerous benefits and soon evolved to include remote monitoring and diagnostics as part of the contract. Starting with a willing customer and a dial-up modem, Siemens Westinghouse installed PC based data acquisition software at a handful of plants to collect data and periodically send it to the monitoring center in Orlando. The data files were in an undocumented binary format, so a vendor-supplied plotting tool was used to export the data as a comma-separated value (CSV) file that was then loaded into a database and analyzed using an “off-the-shelf” spreadsheet program. This analysis was a manual effort and the decision to create a gas turbine rule base was made. GTAID (Gas Turbine Artificial Intelligence Diagnostics) was born. Thus started the rapid expansion of the PCPDS “monitored fleet” and the creation of a department to perform this work called Power Diagnostics®.

GTAID required PCPDS to undergo some major redevelopment. The plan was to monitor a fleet of hundreds or even thousands of gas turbines and generators. This constituted a greater number of plants than had been previously monitored. The first step was to change PCPDS from an online monitoring system based on files to a batch processing system whose inputs and outputs would be stored in a database. A workflow system also needed to be developed to automatically detect the arrival of new data and feed it through the rule base system. Shortly after the first version of database-enabled PCPDS was released a web portal was commissioned to give Monitoring Engineers access to the rule base results.

A Rule Base Engineering Team was formed with the mission of taking the prototype gas turbine rule base and adapting it to analyzing a fleet. To do so, it was necessary to first understand what state the turbine was in: was it on turning gear, or online producing power, or in some stage of a startup or shutdown. This effort was challenging given three factors:

1. the quality of the data received
2. the variation in instrumentation from unit to unit
3. the degree of customization on each unit

Each time a new unit was brought into the monitored fleet, its rule base would have to be adjusted to reduce false positive and negative diagnoses. The rule base logic for each unit would also need to be updated to take advantage of new sensors.

As the rule base grew, managing the variability in the fleet configurations became more and more difficult. As is often the case in such situations, the Rule Base Engineering Team first divided up the rules into groups with each group containing the logic necessary to analyze a component or sub-system. The rule groups representing components were then further divided into types. For example, the combustion sub-system of the Westinghouse-design gas turbines had multiple combustion system designs (e.g. DF-42, DLN) available. Furthermore, an ontology was defined that allowed a rule base engineer the ability to pick “off-the-shelf” rule groups to configure a unit-specific rule base containing only the rules needed to monitor that unit.

A challenge facing rule base engineering that did not come from the sensors or the rule base itself was the fact that the behavior of units operating in the field changed over time. Some of the unit behavior changes were due to upgrades and modifications to improve the equipment efficiency and performance. Other equipment behavior changes were due to changes in control strategies. This illustrated the importance of situational awareness on the part of the Monitoring Engineer and the advantage of coordinating with the rule base engineer.

While the decision had already been made to move away from manual analysis in spreadsheets and create GTAID using the PCPDS expert system shell, not all of the capabilities of the spreadsheet analysis could be replicated. In particular, state estimation models: an implementation of support vector regression to estimate sensor values over time as a function of specific process drivers. Power Diagnostics® personnel worked to develop a Support Vector Machine (SVM) system to create and maintain mathematical models scalable to a large fleet. Importantly, this system, called Power Monitor, could produce residuals – the difference between a measured value and an expected val-
ue – and the residuals generated by this system were used as inputs into the rule base to aid in diagnosing turbines. This greatly expanded the existing pattern recognition capabilities of PCPDS by introducing a powerful, flexible empirical modeling tool to support the analysis of operating data.

A review of Figure 1 explains why this section of the paper was titled “The Growth Years”. During this period of time, the utilization of the PCPDS expert system shell to support the remote monitoring and diagnostics of turbines and generators grew from a mere handful of units to over 1,200.

![Total PCPDS Monitored Fleet (Turbines and Generators)](image)

**Figure 1.**

**PDS Today: PCPDS**

It has been 30 years since the product launch of PDS at Westinghouse. Today, Power Diagnostics® monitors over 1,200 turbines and generators by processing in excess of 5.5 million sensor-hours a day through the PCPDS rule base engine / expert system. More than 16,000 unique rule base outputs have been developed, with 40% of these configured to diagnose operational issues or hardware issues and 60% sensor issues.

PDS has grown much since the 1980s. At least five re-writes of the rule engine itself have been done to accommodate increasing analysis and performance requirements as well as changes in computing hardware and infrastructure requirements. Methods for managing the complexity of rule sets have also been created to scale the rules horizontally to cover the number of unit and plant configurations. Finally, change management systems were introduced to track changes to the rules themselves and to allow Monitoring Engineers to submit change requests and track their progress.

The design of PDS is considered a unique rule-based expert system in its combination of confidence factors, incremental evidence, and parametric rules that modify their outputs based on the quality of the input data. Without this tool, the ability to monitor a fleet of 1,200 plus turbines and generators, to the extent which they are being monitored today, would not be possible without significantly increasing the number of Monitoring Engineers which would basically make the cost of monitoring prohibitive.

One area currently getting additional development attention is the processing speed of the expert system. This is more related to the rule development/test process rather than the production monitoring environment. Improving the speed of PCPDS is expected to shorten the rule development cycle.

Visibility into the rule base is another area currently getting additional development attention. Development in this area is focused on further improving the transparency of the knowledge encoded in the rule base for Monitoring Engineers.

**Disclaimer**

This paper is based mainly on the recollections of the authors. It reflects solely the opinion of the authors and no other party or entity. No representation, whether as to the accuracy, adequacy, usefulness, completeness or otherwise are made.

**References**


