Rule Based Expert System for Error Log Analysis

Omkar Patil, Umesh Chavan

Abstract: ***Humans have been using their domain expertise intelligently and skillfully for making decisions in solving a problem. These decisions are made based on the knowledge that they have acquired through experience and practice over a course of time, which will be lost after the expert's life ends. Hence, this expert knowledge is required to be stored to a database and a machine could be intelligently programmed which could use this knowledge to make decisions, known as an Expert System (ES). This system tries to emulate the decision-making skills of a domain expert by gathering knowledge of the domain experts, storing it to a knowledge base in rule format, and then using those rules to analyze the given data and provides solutions to the problems. These Expert Systems can be utilized to analyze the system log files, find issues logged into those log statements and provide solutions to the errors that are found in those logs.***

Index Terms: Artificial Intelligence, Expert System, Inference Engine, Knowledge base, Log Analysis, Log statements, Rules.

I. INTRODUCTION

Man-made reasoning applied in the space of PC application which attempts to recreate the conduct and basic leadership abilities of a human in a specific area, is known as an Expert System. It plays out its role by acquiring the important information from its information base and deciphering it as indicated by the client's concern. The information in the information base is refreshed by people who are master in those specific spaces and this framework is used by some non-master clients to get data. Medicinal conclusion, bookkeeping, and so forth are some of the regions where these frameworks are broadly utilized.

Expert System (ES), is a domain of Artificial Intelligence [1], a program that tries to mimic the decision-making behavior along with the skills of a human expert. These systems do not have a specific design, flow diagram or an algorithm, each system can have its own structure as designed by the experts, but all expert systems ultimately have rules; results are always computed through information-based process. Knowledge base along with inference engine form the two main components of an Expert System[2]. The knowledge base is a database for the system, stores all the existing information gathered from sources like domain experts or literatures, and consists of rules represented as 'IF <condition> THEN <action>' including facts. The important factor in expert system is large quantity of information gathered from various domain experts in a structured format.

The task of inference engine is to get a query and fire it against rules and determine which rule matched so that an action could be suggested. The expert system makes use of reasoning to perform query and rule matching, which could be either forward or backward reasoning.

1.1. Log Analysis

Log analysis is the process where log files are analyzed to find any issues or discrepancies recorded by the working software or tool. These log files contain statements that are written at runt time by a software, whose purpose is to be used for debugging errors. These log files have their own structure as described by a developer of the software. Hence, any log analysis tool or application can't use these log statements directly. The log statements must be parsed before they are analyzed. This parsing can be done using the statement tokenization or using regular expressions. Once these statements are parsed, then they can be utilized by any log analysis tool. The goal behind Rule Based Expert System for Error Log Analysis is to use this concept where log files are provided to the expert system as an input, which will then parse those log files and extract error statements from those log files which will then be provided to the inference engine that will utilize the store knowledge base and provide solutions to those errors.

1.2. Rule Based Expert System

Rule-based structures (in any case called Expert systems) are the standard based structures, which uses runs as the data depiction for data coded into the system. The implications of rule-set up together system depend prevalently as for expert knowledge structures, which are system that duplicate the considering human expertise in dealing with a data concentrated issue. Instead of addressing data in an authoritative, static way as a ton of things which are substantial, rule-based system address data to the extent a ton of concludes that figures out what to do or what to wrap up in different conditions. A standard based system is a strategy for encoding a human ace's data in a slight area into an automated structure. A standard based structure can be fundamentally made by using a ton of certifications and a ton of concludes that demonstrate the correct conduct on the announcement set. Rules are conveyed as a great deal of if declarations (called IF-THEN principles or creation rules):

The rule states that

IF P, THEN Q

This standard based framework comprises of a couple of fundamental and basic components as follows:

1. An arrangement of facts. These truths are the assertions and should be anything pertinent to the beginning state of the system.

Revised Manuscript Received on July 20, 2020.

***** Correspondence Author

Omkar Patil, Department of Information Technology, Walchand College of Engineering, Sangli, India. E-mail: oapatil24@gmail.com

Umesh Chavan, Department of Information Technology, Walchand College of Engineering, Sangli, India. E-mail: umesh.chavan@walchandsangli.ac.in

2. An arrangement of rules. This contains all moves that should be made inside to demonstrate worthy conduct on the affirmation set. A standard relates the substances in the IF part to some movement in the THEN part. The system should contain simply significant benchmarks and avoid the pointless ones in light of the fact that the number of rules in the structure will impact its display.

1.3. Components of an expert system

- Knowledge base: This information base speaks about actualities and rules. It comprises of information in an area just as rules to take care of an issue, methods and characteristic information important to the space.
- Inference engine: The capacity of this component is to bring the significant information from the information base, decipher it and to discover an answer pertinent to the client's concern. It obtains the guidelines from its information base and applies them to the well-established certainties to gather new realities.
- Knowledge acquisition and learning module: The capacity of this part is to permit the master framework to secure increasingly more information from different sources and store it in the information base.
- User interface: This module makes it feasible for a non-master client to interface with the master framework and discover an answer for the issue.

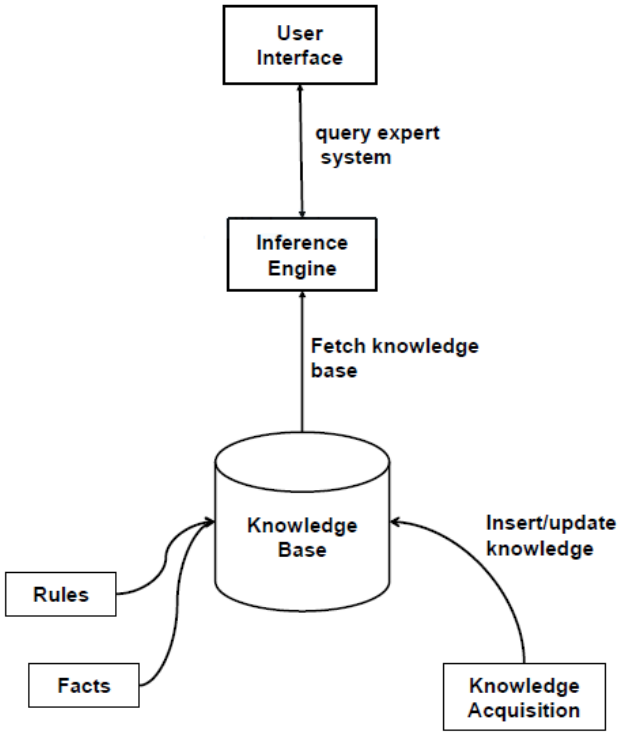


Figure 1. Expert System overview

1.4. Forward Chaining

The inference engine can perform reasoning on the given data using two methods, forward chaining and backward chaining[3]. This work utilizes the method of forward chaining.

This method first tries to prove or satisfy the antecedent part of the rule which then results in evaluation of the consequent part. The opposite of forward chaining is backward chaining [1].

II. LITERATURE REVIEW

This paper by David A. Sanders et al. [4] proposes a rule based expert system that will assist a disabled person in controlling an electrically powered wheelchair. Information regarding the wheelchair surroundings is fetched from ultrasonic sensors, to assist a disabled person riding the wheelchair in avoiding obstacles detected in the path. Rules were represented in 'IF-THEN' form and generated by combining inputs from sources like steering angle, distance to an object on right, distance to an obstacle on left and an angle to the destination. In case, sensor output increases beyond a threshold, then the collision avoidance rules would be activated. Bernardo Canovas-Segura et al. [5] in this paper states that modern Clinical Decision Support System (CDSS) provides healthcare professionals with relevant knowledge for diagnosis decisions. Health institutions from Europe have released a catalogue of rules that will assist microbiologists, carrying out their tests in evaluating the success of an antibiotic. Example of a rule is "If the microorganism belongs to the Entorbacter clocae species, THEN report as resistant to Amoxicillin - clavulanate". In this paper, the author has proposed a methodology which will enable in importing knowledge from an external source into a CDSS based on rule, which will model clinical knowledge using a clump of rules, having the IF-THEN format. When IF condition in the part is met, the system executes actions indicated by THEN part. This method works with reference ontology (REO), consisting of 4 steps, first is to define Reference Ontology, which represents concept, relations along with properties drawn out form sources. Next is mapping of local terms where REO individuals are allowed to be added to the knowledgebase. Then is extending the reasoning properties, where properties of production rule engine are extended with ontology reasoning, which can be performed where rules are simulated by ontology reasoning or by consuming an external ontological reasoning. For implementation purposes, Drools is used which is Java based rule engine.

As stated in this paper by Wei He et al. [6] in WSN, fault is represented by abnormal values, which causes inaccuracy of data fusion. Hence, it's important to research diagnosing faults for data stream in WSN which could find and correct the fault, guaranteeing accuracy of data fusion. Fault diagnosis of WSN involves qualitative knowledge approach wherein knowledge of an expert is utilized in providing rapid diagnosis using symbolic reasoning, whereas in data driven approach, huge data is acquired and fault diagnosing is performed by classifying data features. This paper states that fault diagnosis has two parts: fault detection and determination of fault type. In fault detection, analysis is done on acquired data and sensors with abnormal outputs are found. Then the sensor type and data from sensors is used to get antecedent attributes and the model of determination of fault type is built based on hierarchical Belief Rule Base (BRB). Attributes in antecedent are decided considering time, space and attributes correlation.

The data's antecedent attributes are inputs to the BRB, and corresponding output are inputs to the next layer BRB.

This paper is written by Ahthasham Sajid et al. [3] wherein the author describes Opportunistic network where reason for unreliable communication is the intermittent connectivity between the source and destination. Various factors could cause congestion in links which could be topology, bandwidth, network usage method or hardware failure. In opportunistic networks, asynchronous mode is used for communication. Congestion could occur either at link or storage level, due to which this research suggests an expert system based on rules, which will make nodes smart, enabling congestion detection. This system uses forward chaining method and has 3 submodules which are buffer occupancy, previous contact history and custody transfer module. A network node after connecting to other nodes, checking its buffer occupancy is done to categorize node as less, medium or high congested node. Second, determining if the node was encountered before or not. And third, invoking the custody transfer module which will be responsible to conduct successful custody transfer of packets depending on the rules declared. The proposed system will be able to monitor total packets sent from source and received by the destination, packet transfer between the intermedial nodes and measure each nodes packet loss in opportunistic network. Expert System operates on rules having the structure of "If/Then" rules in which "IF" is the antecedent which is provided as an input to the system, which then states the actions required, represented by the consequent. A rule can have many antecedents combined by conjunctions (AND) or disjunctions (OR). This proposed system contains 5 rules formed using 12 facts, where rules look like. Simulations are performed using Java based ONE simulator, by varying size of buffer from 3, 6, 9 and 12 mb. Test results show delivery ratio increase from 1 to 6 % than MaxPop protocol and comparison for buffer size vs overhead ratio shows improvements as well.

In this paper, Felicidad Aguado et al. [7] presents a situation from the medical domain where explanation have a crucial role and that situation is process of donor –patient matching in an organ transplantation unit. The transplantation unit is expected to follow an objective policy that considers medical parameters which is experimentally supported by the existing records, also, these decisions must be easily reproducible and comprehensible. This paper presents a rule interpreter, web-liver, designed for assisting the medical experts in the donor-patient matching a liver transplantation. The final goal of this tools is providing a rule editor and interpreter that the expert can interactively use to test different policies (set of rules), checking not only their accuracy but also the explanation provided for the obtained decisions. The web liver system is a web application that allows creating, manipulating rules through web interface.

III. METHODOLOGY

We propose a rule based expert system for error log analysis. The system is designed to troubleshoot a log for errors and provide solutions to those errors. The system involves parsing the logs and then submitting those logs to Expert System which will apply the stored domain expert (software developer or QA) knowledge on those parsed logs and the provide solution for the errors.

For a log file to be analyzed, the log bundle which is a compressed bundle of log files, gathered from multiple components is uploaded to the system, which then recursively extracts this log bundle to a temporary location. After which the segregation engine starts its execution, which moves only required files from the extracted temporary folder to an internal location. This internal located folder is then taken by the expert system for preprocessing and analysis.

Before the Expert System could analyze the log files these log files have to be parsed. Where each field form the log must be separated into a structured format, or in some cases only error statements from those log files is required be separated.

After the log file has been parsed, all the parsed data is forwarded to the Expert System for the actual analysis process to begin. Expert System is a software application that simulates the behavior of a domain expert. It consists of a knowledge base and an inference engine. The knowledge base is the database which contains the knowledge [5] gathered from a domain expert, and stored in structured form, of rules, which then can be utilized to analyze the logs and suggest solutions as a domain expert would do. Second component is inference engine whose task is to fetch all the rules form the knowledge base, apply those rules on to the input data. The triggered

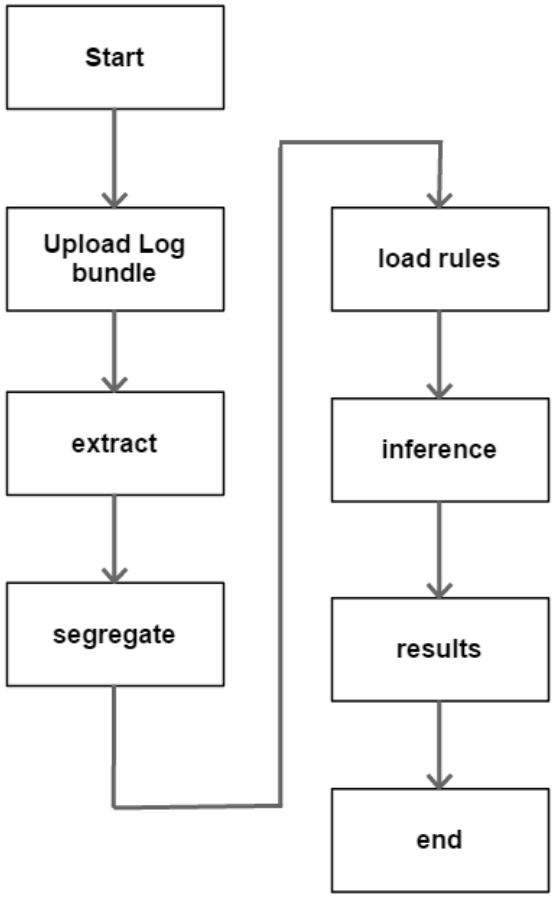


Figure 2. Steps in rule based expert system for error log analysis

rules will provide a solution to the matched log statements or error, for the issue to be resolved.
Here the rules here have a specific format, which is
(Rule Name)
IF
<Component> and <rule statement>
and <tags> and <parameters>
Then
Resolution and knowledge Base links
Each rule will have a rule handler, which will decide how to apply the rule to the log statements. There are multiple log files in the provided log bundles, each having its own format. Hence there's a component name field associated with each rule, which will enable the inference engine to pick only relevant log file associated with that component. After all the rules have been executed, a resulting list of matched or triggered rules along with the associated solution is returned to the user.

IV. RESULTS AND DISCUSSION

1. Choosing the Expert System Development Tool

An Expert System does not have any defined algorithm or programming standard. All it requires is proper representation of the knowledge that it will use for reasoning.

For the development purpose of this Rule based Expert System, we have used python programming language. The user interface and knowledge acquisition system is developed in python using the Django web development framework. And for the purpose of knowledge storage the system uses SQLite database.

2. Knowledge Acquisition

This is one of the most important component of an Expert System, whose purpose is to acquire knowledge from domain expert or gather knowledge [4] from various sources like literature, web pages, etc and store it in a structured format which is then utilized by the inference engine for reasoning.

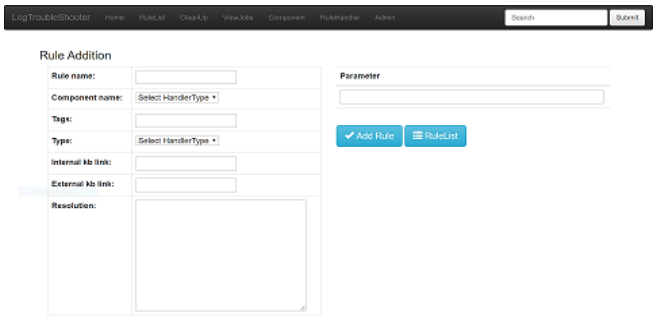


Figure 3. Knowledge Acquisition page

In case of system that is proposed in this paper, the system provides a web portal which has a form to add new rules and knowledge. This form is used by the domain expert or a software developer who has knowledge about the error and the solution corresponding to that error. The rule will contain rule name, error statement, rule handler which will be responsible for how the error statement is processed along with the component for which the rule is written and the solution for the error or the THEN part of the rule along with some extra informational reference links.

3. Production Rules

Production rules are the Business Rules used in this system. These rules are combination of error statements, rule handlers and facts related to the error statement.

A general format of the rule can be given as:

IF
ruleHandler(error_statement)
AND
Component_name
AND
Tags
THEN
Solution and reference links

4. Interface Design

The interface for the expert system is designed using Python and Django. There is a web page for the user or an administrator to upload a log bundle or provide path to the log bundle.

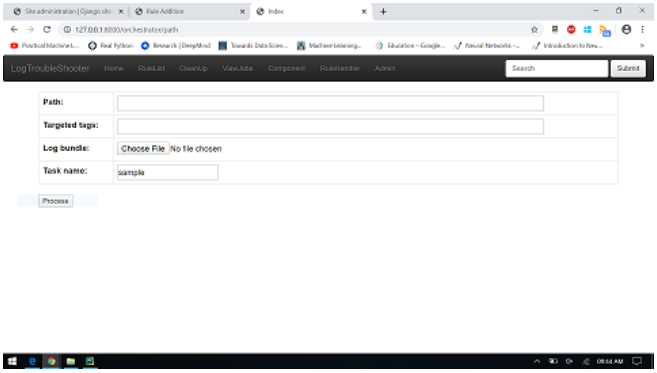


Figure 4. Expert System Homepage

Once a user or administrator uploads the log bundle, he is redirected to a task list page, where all the currently executing tasks list is display where the user could check the progress status of his submitted request.

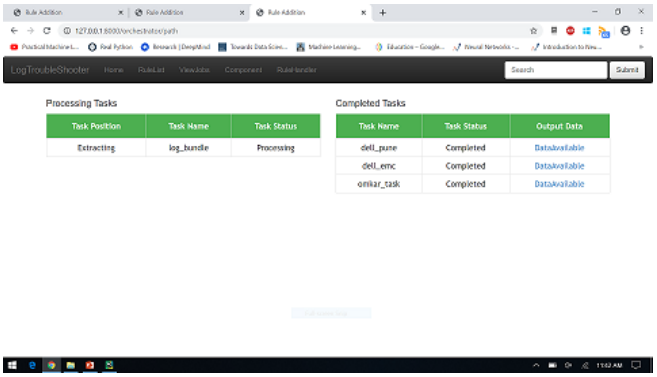
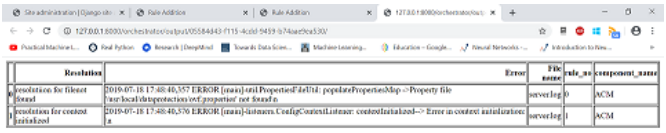


Figure 5. Tasks list

And when the task submitted by the user is completed, a link is activated which will redirect him to the results page where the results will contain the error statement for which a rule has matched, solution for that error along with some extra information like the file in which the error was found and the component for which the error has been found.



Resolution	Error	File name	rule_no	component_name
resolution for fileout found	2019-07-18 17:48:40,357 ERROR [main] util.PropertiesFileUtil: populatePropertiesMap ->Property file '/usr/local/ibm/protection/ext/.properties' not found in	server.log	0	ACM
resolution for context initialized	2019-07-18 17:48:40,570 ERROR [main] listeners.ConfigContextListener: contextInitialized--> Error in context initialization: a	server.log	1	ACM

Figure 6. Results

V. CONCLUSION AND SUGGESTIONS

A. Conclusion

From the results acquired from the above study, it can be concluded that

1. Expert system can be designed to perform log analysis for errors.
2. Expert system can be used to process the error statements and provide solutions to those errors.
3. Latest programing languages can also be used in designing the expert system along with knowledge base and writing the inference engine which will perform the reasoning over the statements.

B. Suggestions

For the further development in this paper, following suggestions are given:

1. Design that could support time stamps for analyzing error statements.
2. Utilizing some advanced schema less database.
3. Providing recommendations related to the most recent frequently occurring errors.

REFERENCES

1. R. Pratama *et al.*, "Expert system for diagnosing vertebrate animals with visual prolog 8.0," *Proc. - 2018 3rd Int. Conf. Inf. Technol. Inf. Syst. Electr. Eng. ICITISEE 2018*, pp. 100–104, 2019.
2. Z. Dong, J. Zhao, J. Duan, M. Wang, and H. Wang, "Research on Agricultural Machinery Fault Diagnosis System Based on Expert System," *Proc. 2018 2nd IEEE Adv. Inf. Manag. Commun. Electron. Autom. Control Conf. IMCEC 2018*, no. Imcec, pp. 2057–2060, 2018.
3. A. Sajid and K. Hussain, "Rule Based (Forward Chaining/Data Driven) Expert System for Node Level Congestion Handling in Opportunistic Network," *Mob. Networks Appl.*, vol. 23, no. 3, pp. 446–455, 2018.
4. D. A. Sanders, A. Gegov, M. Haddad, F. Ikwan, D. Wiltshire, and Y. C. Tan, "A rule-based expert system to decide on direction and speed of a powered wheelchair," *Adv. Intell. Syst. Comput.*, vol. 868, pp. 822–838, 2018.
5. B. Cánovas-Segura, A. Morales, J. M. Juarez, M. Campos, and F. Palacios, "A lightweight acquisition of expert rules for interoperable clinical decision support systems," *Knowledge-Based Syst.*, vol. 167, pp. 98–113, 2019.
6. W. He, P. L. Qiao, Z. J. Zhou, G. Y. Hu, Z. C. Feng, and H. Wei, "A New Belief-Rule-Based Method for Fault Diagnosis of Wireless Sensor Network," *IEEE Access*, vol. 6, pp. 9404–9419, 2018.
7. F. Aguado, P. Cabalar, J. Fandinno, B. Muñiz, G. Pérez, and F. Suárez, "A Rule-Based System for Explainable Donor-Patient Matching in Liver Transplantation," *Electron. Proc. Theor. Comput. Sci.*, vol. 306, pp. 266–272, 2019.
8. E. K. Gebre-Amanuel, F. G. Taddesse, and A. T. Assalif, "Web based expert system for diagnosis of cattle disease," *MEDES 2018 - 10th Int. Conf. Manag. Digit. Ecosyst.*, pp. 66–73, 2018.
9. F. Başçiftçi and E. Avuçlu, "An expert system design to diagnose cancer by using a new method reduced rule base," *Comput. Methods Programs Biomed.*, vol. 157, pp. 113–120, 2018.
10. X. Xu, X. Yan, C. Sheng, C. Yuan, D. Xu, and J. Yang, "A Belief Rule-Based Expert System for Fault Diagnosis of Marine Diesel Engines," *IEEE Trans. Syst. Man, Cybern. Syst.*, pp. 1–17, 2017.
11. M. Peña, F. Biscarri, J. I. Guerrero, I. Monedero, and C. León, "Rule-based system to detect energy efficiency anomalies in smart buildings, a data mining approach," *Expert Syst. Appl.*, vol. 56, pp. 242–255, 2016.
12. J. S. Jadhav, D. K. Nalawade, and D. M. M. Bapat, "Rule-Based Expert System and Its Application with Special Reference to Crimes Against Women," *3rd Int. Conf. Work. Recent Adv. Innov. Eng. ICRAIE 2018*, vol. 2018, no. November, pp. 1–4, 2019

AUTHORS PROFILE

Omkar Patil is a Masters of Technology student at Walchand College of Engineering, Sangli. His field of interest are artificial intelligence, machine learning, deep learning, etc.

Umesh Chavan is a Professor at Department of Informaton Technology, Walchand College of Engineering, Sangli. He is currently pursuing his PhD. His field of interest are artificial intelligence, machine learning, deep learning, computer vision, etc.