FUZZY DIAGNOSIS MODULE BASED ON INTERVAL FUZZY LOGIC: OIL ANALYSIS APPLICATION

Antonio Sala
Systems Engineering and Control Dept., Univ. Politécnica de Valencia
Cno. Vera s/n, 46022 VALENCIA, SPAIN

Bernardo Tormos, Vicente Macián, Emilio Royo
CMT Motores Térmicos, Univ. Politécnica de Valencia
Cno. Vera s/n, 46022 VALENCIA, SPAIN

Keywords: fuzzy expert systems, fuzzy diagnosis, uncertain reasoning, interval fuzzy logic.

Abstract: This paper presents the basic characteristics of a prototype fuzzy expert system for condition monitoring applications, in particular, oil analysis in Diesel engines. The system allows for reasoning under absent or imprecise measurements, providing with an interval-valued diagnostic of the suspected severity of a particular fault. A set of so-called metarules complements the basic fault dictionary for fine tuning, allowing extra functionality.

1 INTRODUCTION

In diagnosis of industrial processes, there is a significant practical interest in developing technologies for a more effective handling of the information available to ease the procedures of inspection and maintenance (I/M) by means of greater automation.

Computer-aided diagnosis is one of the earliest fields of applications of artificial intelligence tools (Russell and Norvig, 2003). Logic and statistical inference have been tried in previous applications.

Indeed, the full diagnostic problem under uncertain data would need to be considered in a probabilistic framework. Taking into account the “gradualness” of the symptoms and possible diagnostics of varying degree of severity (captured by fuzzy logic), the most complete approach would be setting up a continous Bayesian network. This paradigm arose in the last decade as a probabilistic alternative to reasoning, superior to truth-maintenance approaches in some cases (see (Russell and Norvig, 2003) and references therein). However, inference on this paradigm is intractable in a general case (NP-hard). If the amount of uncertainty is low (if a significant subset of the possible measurements is always obtained and the “determinism” of the underlying system is acceptable), then fuzzy logic-based approaches to reasoning may be a viable solution in practice. This is the case of some industrial diagnosis problems, such as oil analysis, to which the system in development is targeted.

Logic uncertainty can be accommodated by possibility theory (Cayrac et al., 1996), or by interval-valued fuzzy logic (Entemann, 2000). The second approach is the one followed in this work. Other works discussing condition monitoring (diagnostic and supervision tasks) using fuzzy logic are, for instance (Carrasco and *et. al.*, 2004; Chang and Chang, 2003). Condition monitoring can also be dealt with with model-based approaches (Isermann and Ballé, 1997), if enough quantitative descriptions of the system are available.

This paper presents the structure of a fuzzy inference module that incorporates some innovations easing the setting up of rules and improving the quality of the final diagnostic conclusions. In particular, the use of interval fuzzy logic, the methodology to deal with exceptions and the possibility of expressing different alternatives for the same diagnostic and, if they do not aggree, firing a fuzzy contradiction warning.

An application of the system is presently being tried on an oil analysis task whose main requirements appear in (Macián et al., 1999).

This paper presents, in two sections the structure of the fuzzy condition monitoring module being developed and the key concepts of the oil analysis application in which the possibilities of the system are being tested.

Sala A., Tormos B., Macián V. and Royo E. (2005).
FUZZY DIAGNOSIS MODULE BASED ON INTERVAL FUZZY LOGIC: OIL ANALYSIS APPLICATION.
In *Proceedings of the Second International Conference on Informatics in Control, Automation and Robotics*, pages 85-90
DOI: 10.5220/0001161200850090

2 THE FUZZY CONDITION MONITORING MODULE

The presented fuzzy condition monitoring module is structured in the following main submodules:

- measurement preprocessing
- fuzzy rule base inference
- postprocessing of the conclusions

2.1 Measurement preprocessing

Raw data from sensors may need some sort of pre-processing prior to rule evaluation. Indeed, if non-linearity inversion, statistical calculations, dynamic processing, etc. is carried out beforehand, then the subsequent rules will be simpler. This preprocessing is, however, application-specific in most cases (see Section 3).

Incomplete or absent measurements: In the case of incomplete information, the measurements can be given in interval form, and fuzzy reasoning will be carried out via generalisation of ordinary rules to interval-valued logic values, as described in Section 2.3, giving rise to an interval output of estimated severities.

2.2 Fuzzy inference module

The fuzzy inference module has been also built with different submodules. It has a “variable definition submodule”, a “rulebase definition submodule” and an “inference engine”.

Variable definition. The name, operating range and applicable ”concepts” (fuzzy sets) on the variables are defined via a suitable syntax. An example appears below on the variable ”CU” (copper concentration):

```
CU NORMAL 5 1     CU HIGH 5 20
CU VERYHIGH 50 150
```

the first number defining the support of the fuzzy set, the second one defining the core. The last line defines, for instance, that the concentration of copper is not ”very high” if it is below 50 ppm, and then, gradually starts to be considered ”very high” up to 150 ppm where it is considered 100% abnormal, indicating that rules related to this concept would fire a ”severe” fault. In the intermediate ranges, the rules would conclude a fault with an ”intermediate” severity.

Rulebase definition. The rulebase is defined by means of a set of rules in the form:

```
Disorder  Symptom-List END
```

They conform the core of the rulebase, and the elements in the symptom list are assumed to be linked by an ”AND” connective. For examples, see Section 3. Inference is carried out by evaluating the minimum of the severity of the symptoms in the symptom list of a particular disorder, and assigning that value to the severity of the associated disorder (see later). If some ”OR” connectives were to be used, it can be done by means of the metarules to be defined.

Symptom relevance modifier. Each symptom may be affected by a coefficient indicating that its presence confirms the fault, but its absence (in the presence of the rest) indicates a milder severity.

Metarules. The basic rules can be used to detect particular situations of interest, with an estimated interval of severity as a result. These situations are, many times, the ultimate faults to be detected.

However, there are occasions where they must be combined. This combination may have a logical interpretation in terms of AND, OR, NOT; in this case, the so-called metarules are introduced to handle the situation. One possible structure is:

```
DISORDER IF LOGIC-EXPRESSION
```

where `LOGIC-EXPRESSION` is any user-defined combination of symptoms or previously inferred atomic disorders, with conjunction, disjunction and negation operators. This rules will be denoted as `MIF` rules.

Note that the basic rules could have been embedded into this syntax. However, the proposed one allows for the incorporation of relevance modifiers that might be more cumbersome in the middle of a logic expression.

Another type of metarule is the one in the form:

```
DISCARD Disorder IF LOGIC-EXPRESSION
```

used, for instance, to discard a ”general” fault if a more specific situation sharing the same symptoms (plus some other ones) is encountered, or to set up “exceptions” to rules. Again, in Section 3, some examples appear. Other types of metarules (UNION and ALTERNATIVE) will be later described.

2.3 Inference Methodology

This section discusses the inference mechanism used to evaluate the presented rules and metarules. The core of the inference system works with fuzzy *interval uncertain propositions*. In these propositions, their truth value is an *interval* $[\nu, \pi]$, where $0 \leq \nu \leq 1$ and $0 \leq \pi \leq 1$. It describes partial knowledge: the minimum value it can attain given the present information is called *necessity* ν and the maximum value will be denoted as *possibility* π. If they do coincide, then the proposition can be considered an ordinary fuzzy proposition. The interpretation of possibility and necessity is related but not equivalent to other approaches such as (Cayrac et al., 1996).

Connectives. Connectives allow combination of atomic propositions into more complex ones by conjunction, disjunction and negation. Assuming a particular connective for a fuzzy proposition is given (such as T-norms for AND, T-conorms for OR, etc. (Weber, 1983) in a non-intervalic fuzzy logic) in the fuzzy-uncertain framework, the resulting interval will be evaluated with interval arithmetic, *i.e.*, given a fuzzy connective $C : [0,1] \times [0,1] \to [0,1]$, $Y = C(X_1, X_2)$, its interval version is[1]:

$$[\nu_Y, \pi_Y] = \left[\min_{\substack{\mu_1 \in [\nu_1, \pi_1] \\ \mu_2 \in [\nu_2, \pi_2]}} C(\mu_1, \mu_2), \max_{\substack{\mu_1 \in [\nu_1, \pi_1] \\ \mu_2 \in [\nu_2, \pi_2]}} C(\mu_1, \mu_2) \right] \tag{1}$$

Regarding negation, the truth degree of $\neg p_1$ is defined as the interval $[\nu = 1 - \pi_1,\ \pi = 1 - \nu_1]$.

For instance, using the minimum and maximum as conjunction and disjunction operators, the intervalisation of AND and OR operations is:

$$[\nu_1, \pi_1] \wedge [\nu_2, \pi_2] = [\min(\nu_1, \nu_2), \min(\pi_1, \pi_2)] \tag{2}$$
$$[\nu_1, \pi_1] \vee [\nu_2, \pi_2] = [\max(\nu_1, \nu_2), \max(\pi_1, \pi_2)] \tag{3}$$

Let us consider now how inference is carried out in the basic rules and in the metarules.

Basic Rules: The AND intervalic operation (2) is used (or its trivial generalisation to more intervals). Furthermore, if a particular symptom has an "irrelevance factor" ρ its membership value μ is transformed to $\rho * (1 - \mu) + \mu$ before carrying out the interval conjunction.

In fact, the implemented approach considers the so-called inference error (Sala and Albertos, 2001) so that given a logic value μ, the inference error is $e = 1 - \mu$ (extended to interval arithmetic). Then, given a list of q symptoms in a rule, the overall inference error is:

$$E = \sqrt[p]{\sum_{i=1}^{q} (e_i)^p} \tag{4}$$

So with $p = 2$ the methodology could be denoted as Euclidean inference. With $p \to \infty$, the result is the same as the interval AND (using minimum) described above. By fixing the value of p the user may specify a different behaviour (the lower p is, the more severity is substracted due to partially non-fired symptoms). The formula (4) can be generalised to intervals, obtaining the lowest inference error by using the norm of the minimum inference error of each symptom, and the highest one with the maxima of the inference error intervals. Membership values are recovered from the resulting inference error figures by means of a negation formula.

Metarules: The logic operations will use the above intervalar expressions when evaluating any LOGIC-EXPRESSION in `MIF` metarules.

In the `DISCARD` metarules, the interval-arithmetic substraction will be used, *i.e.*:

`DISCARD` $[\nu_1, \pi_1]$ `IF` $[\nu_2, \pi_2]$

will give as a result the interval $[\nu_1', \pi_1']$:

$$\nu_1' = \max(0, \nu_1 - \pi_2) \qquad \pi_1' = \max(0, \pi_1 - \nu_2)$$

Membership value transformations. In some cases, one would like to introduce a set of rules detecting conditions for an intermediate fault and different conditions for a severe one (say, F1):

IF COND1 THEN F1 INTERMEDIATE
IF COND2 THEN F1 SEVERE (5)

To deal with this case, the tool under discussion allows *linear* transformations of membership via the so-called `UNION` metarule:

```
UNION FAULTNAME
  IDENT1 l11 l12 IDENT2 l21 l22 ...
```

The coefficients l_{i1} and l_{i2} define a linear transformation $\mu = l_{i1} * (1 - \mu) + l_{i2} * \mu$ to be carried out on the membership of identifier "IDENT-i". Afterwards, an interval-logic OR is carried out. Obviously, to combine a particular condition with no membership transformation, the setting $l_{i1} = 0$ and $l_{i2} = 1$ must be used[2]. For instance, the above case (5) would be encoded by:

```
UNION F1
   COND1 0 0.5 COND2 0.5 1
```

[1] Those expression apply when the propositions refer to *independent, non-interactive* variables. As in ordinary interval arithmetic, the above expression (1) in complex propositions yields conservative (excessively uncertain) results with correlation and multi-incidence in the arguments.

[2] The linear mapping actually implemented is not exactly the one above. It is:

$$\mu' = \begin{cases} l_{i1} * (1 - \mu) + l_{i2} * \mu & \mu \geq 0.02 \\ 0 & \mu < 0.02 \end{cases} \tag{6}$$

In that way, it can be specified that an intermediate severity fault must be suspected if any nonzero activation of a particular condition holds, but no firing will occur if none of the conditions are active above a significant threshold (0.02).

Alternatives. A closely related situation arises when several alternatives for diagnosing the same fault exist. If all measurements were available, all of them should produce the same result so specifying only one of them in the rulebase will do. However, to improve results accounting for missing or imprecise measurements, several of these alternative rules may be intentionally specified. In that case, to allow combining different alternatives into one diagnostic, the *intersection* of the intervals produced by inference on each of them will be the produced conclusion of the inference.

This is implemented in the current tool by an `ALT` metarule, with a syntax similar to the union metarule:

```
ALT FAULTNAME
  IDENTIFIER1 l11 l12 IDENTIFIER2 l21 l22 ...
```

allowing also a linear membership transformation (6) before the interval intersection is calculated.

For instance, let us assume that, after the membership transformations, if any, alternative A yields $[\nu_1, \pi_1]$ and alternative B yields $[\nu_2, \pi_2]$ as estimated severity intervals. If $\pi_1 > \nu_2$, then the intersection is not empty and the following estimated severity interval is produced:

$$[\max(\nu_1, \nu_2), \min(\pi_1, \pi_2)]$$

Otherwise, the system outputs the interval $[\pi_1, \nu_2]$ flagged with a *contradiction warning*, as the intersection is *empty*. If $\nu_2 - \pi_1$ is small, then the contradiction level is small and the above interval can be accepted as an orientative result. If it is a large number, it means that different alternatives give totally different results so an error in the rulebase definition or a fault in one of the measurement devices providing the data must be suspected.

The results of the inference is a list of truth values of the disorders. Those truth values are to be interpreted as the "severities" (from incipient to severe) of the associated disorders.

Post-processing. The output of the expert system (interval of estimated severity) is translated onto a summarised statement. Each value of severity is mapped to a linguistic tag "negligible, incipient, medium, severe", defining an interval of severities for each tag partitioning the full [0,1] range. If both extreme severities of the conclusions have the same tag, then a conclusion in the form:

Fault FAULTNAME is TAG (MINIMUM,MAXIMUM severity)

is extracted. Otherwise, the produced sentence is:

Fault FAULTNAME severity might range from TAG(MIN severity) to TAG(MAX severity)

3 OIL ANALYSIS APPLICATION

Oil analysis is a key technology in predictive maintenance of industrial Diesel engines. Indeed, by determining the amount of wear particles, the composition of them, and other chemicals in the oil, a sensible set of rules can be cast to allow a reasonably accurate prediction of the oil condition and/or some likely engine malfunction.

Expert systems based on binary logic have been developed for the application (Macián et al., 2000), but the use of a fuzzy logic inference engine is considered advantageous and it is being evaluated.

An application of the above ideas is under development at this moment. Let us discuss some issues on its development.

Preprocessing. When acquiring information from a particular engine, some observations have the same meaning for all engines to be diagnosed. However, other ones need the use of historical data to generate "normalised" deviations taking into account statistical information for a particular engine *brand* or *model*, or a particular *unit* with special characteristics. In this way, the rulebase conception can be more general (applied without modification to a larger number of cases) if the data are suitably scaled and displaced prior to inference or, equivalently, fuzzy sets are modified according to the particular engine being diagnosed.

In some measurements, the procedure involves normalising the deviation from the mean expressing it in variance units, and generating an adimensional quantity. The statistical data are calculated from a database of previous analysis, classified in brand, model (and also from historical records from the same engine).

Other variables are transformed to "engineering" units, having a more suitable interpretation than the raw readings (for instance, oil viscosities are expressed as a percentage of a reference value from fresh oil characteristics, instead of the centiStoke measurement).

Also, in order to consider real engine behaviour, oil consumption and fresh oil additions are considered leading to obtain a *compensated* wear element concentration more representative of engine status.

Knowledge base. At this moment, the team is in process of acquiring and refining a knowledge base with diagnosis rules.

The basics of the knowledge to be incorporated on the expert system lie on the following facts.

System is focused on automotive engines diagnosis (trucks, buses and general and road construction equipment), and so, the different parameters to measure are (Macián et al., 1999):

- Oil properties: viscosity, Total Base Number (TBN) and detergency.
- Oil contaminants: Insoluble compounds, fuel dilution, soot, ingested dust (silicon), water and glycol.
- Metallic elements: iron, copper, lead, chrome, aluminum, tin, nickel, sodium and boron.

Other measurements could be performed upon the oil sample, but with these basic parameters a good diagnosis can be achieved. Systems developed for other types of engines could choose other parameters taking into account the particularities of these types of engines.

Let us consider, as an example, the kind of knowledge involved on the dust contamination detection.

Silica and silicates are present at high concentrations in natural soils and dusts. It is for this reason that silicon is used as the most important indicator of dust entry into an engine. There have been several studies done on the causes of premature wear in components and results vary from study to study but one thing is clear: external contamination of lube oil by silicon is a major cause of accelerated wear. Particles of airborne dust vary in size, shape and abrasive properties and in an engine the ingress of atmospheric dust takes place primarily through the air intake. Those particles, not retained by filters, and similar in size to the oil film clearance in main lubricated parts of the engine do the maximum damage. Once the dust particle has entered an oil film it forms a direct link between the two surfaces, nullifying the effect of the oil film, thus, the immediate consequence is a ”scratching” of the surface as the particle is dragged and rolled across the surfaces. The second and potentially more serious problem is that once the dust particle is introduced in between the two surfaces, it changes the loading of the surface from an even distribution to a load concentrated on the particle with a huge increase in pressure at this point. The increase in pressure causes a deflection of the surface, which will eventually result in metal fatigue and the surface breaking up. As soon as a dust entry problem occurs there is an increase in the silicon concentration into the oil and an acceleration of the wear pattern. As long as the oil samples are being taken at regular intervals in the correct manner, the dust entry will be detected at a very early stage. If an effective corrective action is taken, the life of the component will be significantly increased, reducing maintenance costs.

Diagnosis of silicon contamination (dust ingestion):

- If normal wear patterns combine with high silicon readings in oil analysis, there are three possibilities: a silicone sealant, grease or additive is in use mixing with engine oil, an accidental contamination of the sample has occurred or dust ingestion is in the first stage and no others wear patterns are present yet (too lucky situation). Action recommend to the maintenance technicians must be to check if an additive, grease or sealant has been used recently on the engine and make sure that the correct sampling technique was used. An inspection of the air admission system on the engine will be necessary if previous action results negative.
- Increased engine top-end wear (iron, chromium, or aluminium concentration rises up). This increased engine top-end wear is caused by airborne dust that has been drawn into the combustion chamber being forced down between the ring, piston and cylinder. Dust origin is caused by a defective air cleaner or a damaged induction system. Actions to be taken by maintenance technicians are inspect the air filter element thoroughly, and check its seals and support frame for damage and distortion and check too the pleats for damage. If there is any doubt about a filter element, it should always be changed. If the leak was found, it is necessary to repair the leak and determine the condition of the engine by checking compression or blowby.
- Increased engine bottom-end wear (lead, tin or copper concentration rises up). This situation indicates that dirt is basically getting into the lube oil directly and not past the pistons and rings. The likely sources are: leaking seals, defective breather, damaged seal on oil filler cap or dipstick, or dirty storage containers and/or top-up containers. Recommended action to be taken by technicians must be that any dust that is in the oil will be pumped through the oil filter before entering the bearings. Therefore, the first step is to examine the oil filter looking for dust contamination or bearing material. If excessive dust is found, thoroughly check all seals and breathers, etc. Check the oil storage containers and top-up containers for finding the source of contamination.

In the syntax of implemented prototype tool, the rules are:

```
CONTS1 SI NOT NORMAL END
WEAR_1 IF
 FE NOT NORMAL or CR NOT NORMAL
 or AL NOT NORMAL
WEAR_2 IF
 PB NOT NORMAL or SN NOT NORMAL
  or CU NOT NORMAL

 CONTS2 IF CONTS1 and  WEAR_1
 CONTS3 IF CONTS1 and  WEAR_2
 SILICON_CONTAMINATION IF CONTS1
 DUST_INGESTION IF CONTS2 or CONTS3
```

As another example, water problems can be divided into two different sources: an external water contamination or refrigerant leakage, in each case a different behaviour is presented. Additionally, water can evaporated and not to be present in oil. For this case, other fingerprints, that remain in oil when water evaporated must be found, such as: sodium (NA), boron (BO), its ratio (NABO) or glycol (GLIC) (its absence would fire rule CONTW3 at most 70%).

Finally, to take into account a specific situation such as a refrigerant leakage with greats amounts of copper from tube wear caused by water surface corrosion, an specific rule is defined too (CONTW4). In the syntax of the implemented tool, the rules are:

```
CONTW1  WATER NOT NORMAL END
CONTW2  GLIC NOT NORMAL END
CONTW4 CU VERYHIGH END

CONTW3
  NA NOT NORMAL
  BO NOT NORMAL
  NABO ABNORMAL
  GLYC NOT NORMAL 0.7
END

WATERFAULT IF
 CONTW1 or CONTW2 or CONTW3 or CONTW4

DISCARD WEAR IF CONTW4
```

So, based on the ideas from the above rules, a full rulebase is being built at this moment.

4 CONCLUSIONS

This paper presents a prototype fuzzy expert system for oil diagnosis. The flexibility of this system is greater than those of binary rules, allowing for gradation of diagnosis. Also, refinements such as interval-valued memberships, membership transformations, exceptions (discarding), unions and alternatives are included. In this way, the diagnostic capabilities and the readability of the rule base improve substantially.

The oil analysis application in consideration provides a quite complete set of measurements from which expert rules can be asserted with a reasonable reliability. However, for suitable diagnosis on a particular engine, a pre-processing module is essential: this module incorporates records of similar engines (same brand, model, history of the one being monitored, fresh oil characteristics, analytical calculations, *etc.*) so that the fuzzy set definitions are adapted for each case.

The full system is, at this moment, in development and prototype testing stage (comparing with human experts' conclusions and those from preexisting *ad hoc* software based on binary logic), but its preliminary results are promising.

ACKNOWLEDGEMENTS

The research work presented in this paper is financed by project FEDER/MEC DPI-2004-07332-C02-02 (Spain).

REFERENCES

Carrasco, E. and *et. al.*, J. R. (2004). Diagnosis of acidification states in an anaerobic wastewater treatment plant using a fuzzy-based expert system. *Control Engineering Practice*, 12(1):59–64.

Cayrac, D., Dubois, D., and Prade, H. (1996). Handling uncertainty with possibility theory and fuzzy sets in a satellite fault diagnosis application. *IEEE Trans. on Fuzzy Systems*, 4(3):251–269.

Chang, S.-Y. and Chang, C.-T. (2003). A fuzzy-logic based fault diagnosis strategy for process control loops. *Chemical Engineering Science*, 58(15):3395–3411.

Entemann, C. (2000). A fuzzy logic with interval truth values. *Fuzzy Sets and Systems*, 113:161–183.

Isermann, R. and Ballé, P. (1997). Trends in the application of model-based fault detection and diagnosis of technical processes. *Control Engineering Practice*, 5(5):709–719.

Macián, V., Lerma, M., and Tormos, B. (1999). Oil analysis evaluation for an engines fault diagnosis system. *SAE Papers*, 1999-01-1515.

Macián, V., Tormos, B., and Lerma, M. (2000). Knowledge based systems for predictive maintenance of diesel engines. In *Proc. Euromaintenance Conf.*, volume 1, pages 49–54. Swedish Maintenance Society-ENFMS.

Russell, S. and Norvig, P. (2003). *Artificial Intelligence: a modern approach*. Prentice-Hall, 2nd edition.

Sala, A. and Albertos, P. (2001). Inference error minimisation: Fuzzy modelling of ambiguous functions. *Fuzzy Sets and Systems*, 121(1):95–111.

Weber, S. (1983). A general concept of fuzzy connectives, negations and implications based on T-norms and T-conorms. *Fuzzy Sets and Systems*, 11:115–134.