A Fuzzy Symbolic Representation for Intelligent Reservoir Well Logs Interpretation

Tina Yu[1] and Dave Wilkinson[2]

[1] Department of Computer Science
Memorial University of Newfoundland
St. John's, NL A1B 3X5 Canada
tinayu@cs.mun.ca
http://www.cs.mun.ca/~tinayu

[2] Chevron Energy Technology Company
San Ramon, CA 94583 USA
davidawilkinson@chevron.com

Abstract. Well log data are routinely used for stratigraphic interpretation of the earth's subsurface. This paper presents an automatic blocking scheme that transforms numerical well log data into a fuzzy symbolic representation. This representation maintains the character of the original log curves, which is essential for the stratigraphic interpretation, while making the interpretation task easier. Additionally, fuzzy symbols allow effective interpretation under uncertainty embedded in the data set and our knowledge of the earth's subsurface. We present the developed technique and test it on two sets of well logs collected from oil fields in offshore West Africa. The results give sensible well log blocking and resemble the original log curves reasonably well. Based on this fuzzy symbolic representation, an intelligent well logs interpretation system has been developed.

1 Introduction

In reservoir characterization, well log data are frequently used to interpret physical rock properties such as lithology, porosity, pore geometry, depositional facies and permeability [3, 6]. These properties are keys to the understanding of an oil reservoir and can help determining hydrocarbon reserves and reservoir producibility. Based on the information, decisions of where to complete a well, how to stimulate a field, and where to drill next, can be made to maximize profit and minimize risk.

Well log data, ranging from conventional logs, such as *spontaneous potential*, *gamma ray*, and *resistivity*, to more advanced logging technology, such as *Nuclear Magnetic Resonance* (NMR) logs, are sequence of curves indicating the properties of layers within the earth's subsurface. Figure 1 gives an example of *grammar ray*, *neutron* and *spontaneous potential* (SP) logs. The interpreted lithology is listed on the left-hand side.

Well log interpretation is a time-consuming process, since many different types of logs from many different wells need to be processed simultaneously. This paper

O. Castillo et al. (Eds.): Soft Computing for Hybrid Intel. Systems, SCI 154, pp. 417–426, 2008.
springerlink.com

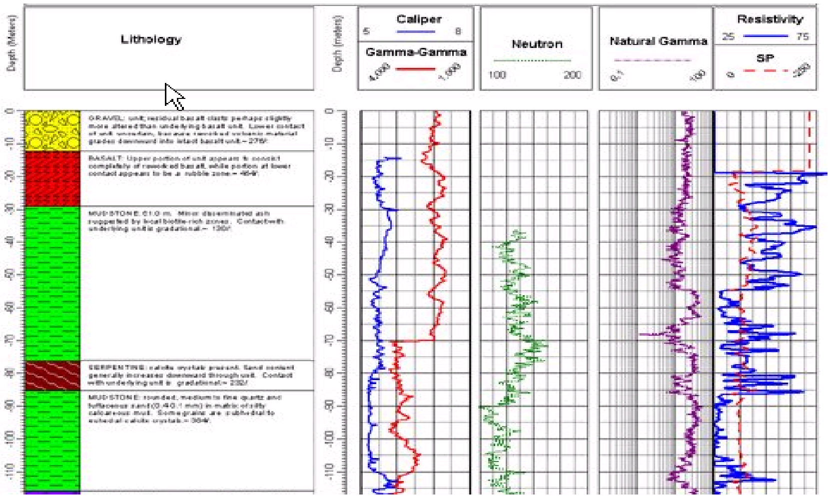

Fig. 1. An example of gamma ray, neutron and spontaneous potential logs. The interpreted lithology is listed on the left-hand side.

presents a technique to automatically transform well log data into fuzzy symbols which maintain the character of the original log curves. This simplified representation not only makes the interpretation task easier but also allows efficient interpretation under the uncertainty embedded in these data sets. The symbolic representation also has advantages over its numerical counter-part in that it is easier for computers to manipulate and process. Based on this representation, we have developed a computer system to perform automatic well-log interpretation. That work is reported in [7].

The paper is organized as follows. Section 2 presents the methodology to transform well log data into a fuzzy symbolic representation. In Section 3, we apply the method to two sets of well log data collected from an oil field in offshore West Africa and show the results. The intelligent well-log interpretation system is briefly described in Section 4. Finally Section 5 concludes the paper.

2 Methodology

The fuzzy symbolic representation is an approximation of well-logs that maintains the trend in the original data. The transformation process has four steps:

- Segmentation of the numerical well log data;
- Determine the optimal number of segmentation;
- Assign symbols to each segment;
- Symbols fuzzification.

These four steps are explained in the following subsections.

2.1 Well Log Segmentation

Well log segmentation involves partitioning log data into segments and using the mean value of the data points falling within the segment to represent the original data. In order to accurately represent the original data, each segment is allowed to have arbitrary length. In this way, areas where data points have low variation will be represented by a single segment while areas where data points have high variation will have many segments.

The segmentation process starts by having one data point in each segment. That is the number of segments is the same as the number of original data points. Step-by-step, neighboring segments (data points) are gradually combined to reduce the number of segments. This process stops when the number of segments reaches the predetermined number of segment.

At each step, the segments whose merging will lead to the least increase in *error* are combined. The *error* of each segment is defined as:

$$error_a = \sum_{i=1}^{n}(d_i - \mu_a)^2$$

where n is the number of data points in segment a, μ_a is the mean of segment a, d_i is the ith data point value in segment a.

This approach is similar to the Adaptive Piecewise Constant Approximation proposed by Keogh et. al. [4]. However, our method has an extra component that dynamically determines the number of segments (see Section 2.2). Another similar work using a different approach to determine the number of segments is reported in [1].

Figure 2 is an example of a well log with 189 data points, which are partitioned into 10 segments. The same data are partitioned into 20 segments in Figure 3. The average value of the data points within each segment is used to represent the original data.

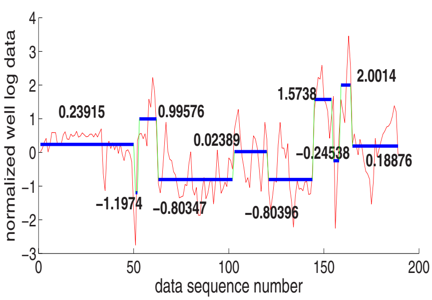

Fig. 2. 10 segments

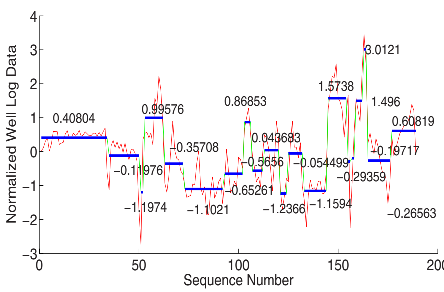

Fig. 3. 20 segments

2.2 Number of Segments

Although a larger number of segments capture the data trend better, it is also more difficult to interpret. Ideally, we want to use the smallest possible number of segments to capture the trend of the log data. Unfortunately, these two objectives are in conflict: the total *error* of all segments monotonically increases as the number of segments decreases (see Figure 4). We therefore devised a compromised solution where a penalty is paid for increasing the number of segments. The new *error* criterion is now defined as the previous total error *plus* the number of segments:

$$f = N + \sum_{i=1}^{N} error_i \qquad \text{where } N \text{ is the number of segment.}$$

During the segmentation process, the above f function is evaluated at each step. As long as this value f is decreasing, the system continues to merge segments. Once f starts increasing, it indicates that farther reducing the number of segments will sacrifice log character, hence the segmentation process terminates. For the log in Figure 2, 50 is the optimal number of segments for the 189 data points (see Figure 5).

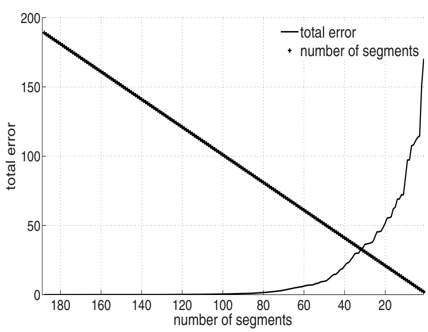

Fig. 4. Number of segments vs. total error

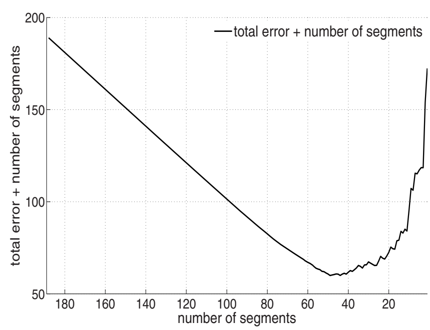

Fig. 5. A compromised solution

2.3 Symbols Assignment

Segmented well logs are represented as a set of numerical values, $\overline{WL} = \overline{s_1}, \overline{s_2}, \overline{s_3} \ldots$, where $\overline{s_i}$ is the mean value of the data within the ith segment. This numerical representation is farther simplified using symbols. Unlike numerical values, which are continues, symbols are discrete and bounded. This makes it easy for any subsequent computer interpretation scheme.

While converting the numerical values into symbols, it is desirable to produce symbols with equal-probability [2]. This is easily achieved since normalized sequence data have a Gaussian distribution [5]. We therefore applied z-transform to normalize the data and then determine the breakpoints that will produce n

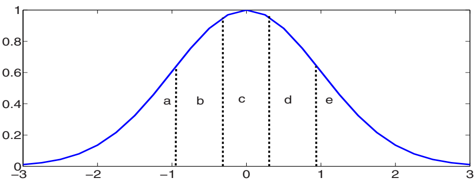

Fig. 6. Using 4 breakpoints to produce 5 symbols with equal probability

equal-sized areas under Gaussian curve. Figure 6 gives the four breakpoints -0.84, -0.25, 0.25 and 0.84 that produce 5 symbols, a, b, c, d, e, with equal probability. If only 3 symbols (a, b and$c$) are used, the breakpoints are -0.43 and 0.43.

Once the number of symbols, hence the breakpoints have been decided, we assign symbols to each segment of the well logs in the following manner: All segments have mean values that are below the smallest breakpoint are mapped to the symbol a; all segments have mean values that are greater than or equal to the smallest breakpoint and less than the second smallest breakpoint are mapped to the symbol b and so on. Figure 7 gives a well log that is transformed using 5 symbols.

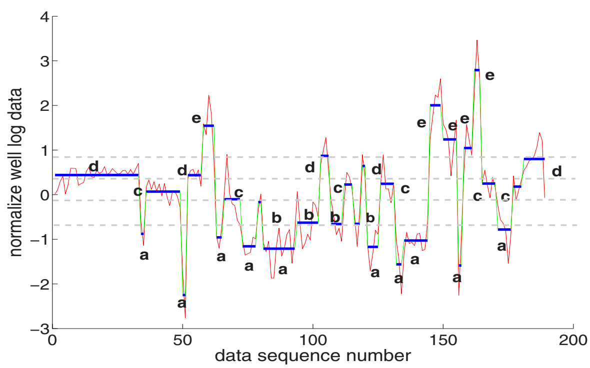

Fig. 7. A well log transformed using 5 symbols

2.4 Symbols Fuzzification

While some segments are clearly within the boundary of a particular symbol region, others may not have such clear cut. For example, in Figure 7, there are 3 segments lie on the borderline of a and b regions. A crisp symbol, either a or b, does not represent its true value. In contrast, fuzzy symbols use membership function to express the segment can be interpreted as symbol a and b with some possibility.

As an example, with crisp symbol approach, a segment with mean -0.9 is assigned with symbol a with 100% possibility (see Figure 8). Using fuzzy symbols designed by trapezoidal-shaped membership functions, the segment is assigned

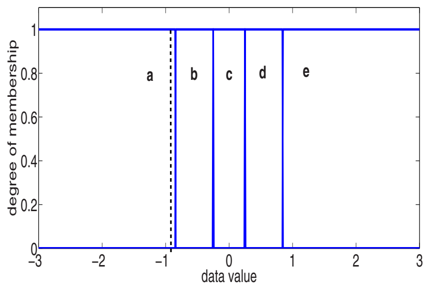

Fig. 8. The data value of -0.9 is transformed as a crisp symbol a

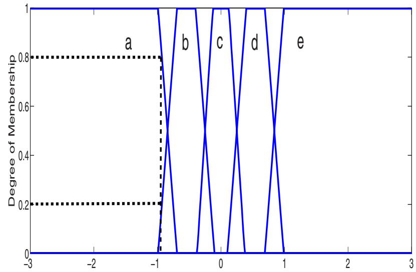

Fig. 9. The data value -0.9 is transformed as fuzzy symbol a (80%) and b (20%)

with symbol a with 80% possibility and symbol b with 20% possibility (see Figure 9). Fuzzy symbol representation is more expressive in this case.

In fuzzy logic, a membership function (MF) defines how each point in the input space is mapped into a membership value (or degree of membership) between 0 and 1. The input space consists of all possible input values. In our case, normalized well log data have open-ended boundaries with mean 0. Since 5 symbols are used to represent a well-log, we need to design 5 membership functions, one for each of the 5 symbols. Additionally, we used 3 symbols to represent core permeability. Three membership functions were also designed for these 3 symbols.

To design a trapezoidal-shaped membership function, 4 parameters are required: f_1 and f_2 are used to locate the "feet" of the trapezoid and s_1 and s_2 are used to locate the "shoulders" (see Figure 10). These four parameters are designed in the following way.

Let c_1 and c_2 be the breakpoints that define symbol n and $c_2 > c_1$:

$$d = \frac{c_2 - c_1}{4}$$
$$f_1 = c_1 - d; s_1 = c_1 + d; s_2 = c_2 - d; f_2 = c_2 + d$$

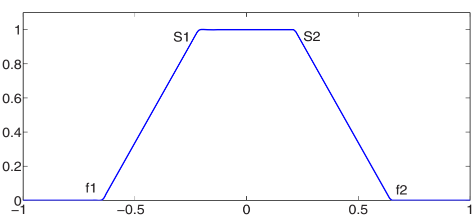

Fig. 10. 4 parameters, f1, f2, s1, s2, define a trapezoidal-shaped membership function

Table 1. Parameters used to design the trapezoidal-shaped membership function for each symbol

data	symbol	f_1	s_1	s_2	f_2
well-log	a	-3	-3	-0.9875	-0.6925
	b	-0.9875	-0.6925	-0.3975	-0.1025
	c	-0.375	-0.125	0.125	0.375
	d	0.1025	0.3975	0.6925	0.9875
	e	0.6925	0.9875	3	3
permeability	a	-3	-3	-0.645	-0.215
	b	-0.645	-0.215	0.215	0.645
	c	0.215	0.645	3	3

There are two exceptions: symbol a has $f_1 = c_1$ and symbol e has $f_2 = c_2$. Table 1 gives the four parameters used to design the membership functions for each symbol.

Once the 4 parameters are decided, the membership function f is defined as follows:

$$f(x, f_1, f_2, s_1, s_2) = \begin{cases} 0, & \text{if } x \leq f_1 \\ \frac{x-f_1}{s_1-f_1}, & \text{if } f_1 \leq x \leq s_1 \\ 1, & \text{if } s_1 \leq x \leq s_2 \\ \frac{f_2-x}{f_2-s_2}, & \text{if } s_2 \leq x \leq f_2 \\ 0, & \text{if } f_2 \leq x \end{cases}$$

Using the described fuzzy symbol scheme, 10 segments lie between two symbol regions in Figure 7 were mapped into fuzzy symbols shown in Figure 11.

In most cases, a reservoir well has multiple logs. To carry out the described transformation process, a reference log is first selected for segmentation. The result is then used to segment the other logs in the same well. After that, fuzzy symbols are assigned to each segmented data.

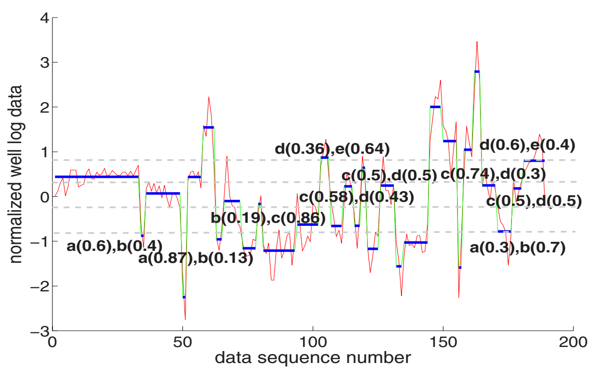

Fig. 11. A well log represented with fuzzy symbols

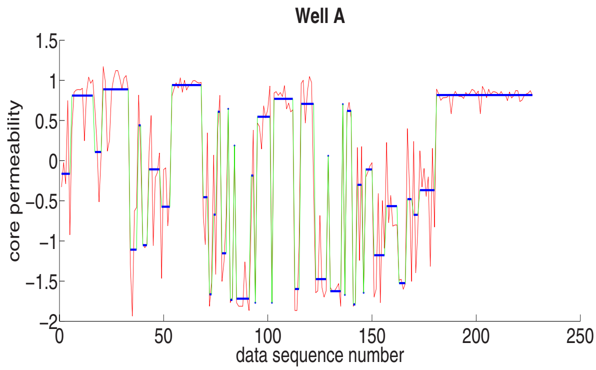

Fig. 12. The transformed core permeability (k)

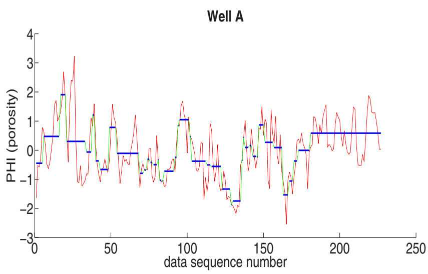

Fig. 13. Transformed PHI log

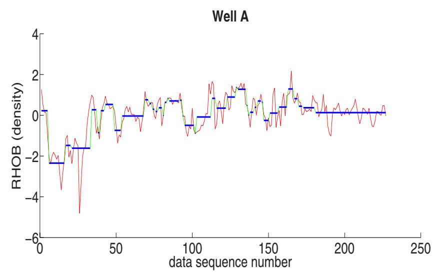

Fig. 14. Transformed RHOB log

3 Testing Results

We tested the developed transformation method on 2 sets of well log data collected from an offshore West Africa field. The first set is from Well A and contains 227 data points while the second set is from Well B and contains 113 data points. Each well has 3 different logs: PHI (porosity), $RhoB$ (density) and DT (sonic log). Additionally, V-shale (Volume of shale) information has been calculated previously [6]. We also have the corresponding core permeability data for these two wells.

In this case, permeability is the interpreted target. It is therefore chosen as the reference log to perform segmentation described in Section 2.1 and 2.2. Based on the segmentation results, the other 4 logs were segmented.

Figures 12, 13, 14, 15 and 16 give the segmented results for Well A. As shown, the results give sensible blocking and resemble the original log curves reasonably well. We do not show the transformed fuzzy symbols on these Figures as they take too much space. Also, due to space constraint, the results of Well B, which have a similar pattern, are not shown here.

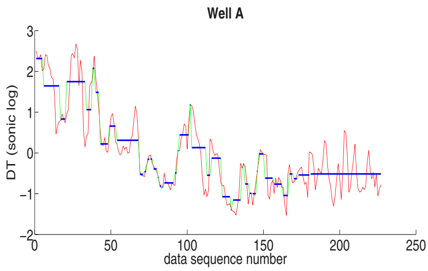

Fig. 15. Transformed DT log

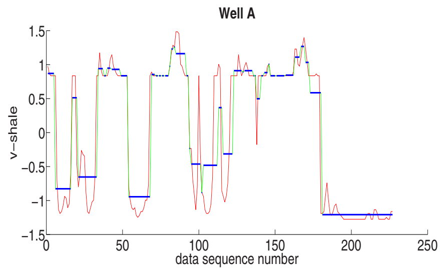

Fig. 16. Transformed V-shale log

4 Automatic Well Log Interpreter

Based on the fuzzy symbolic representation, we have developed an interpretation system that processes the fuzzy symbols and automatically interpret the permeability ranges that are associated with each well log segment. The system applies a co-evolutionary mechanism to evolve fuzzy rule sets. The fuzzy rule set is composed of two fuzzy rules, one classifies high-permeability log segments while the other identifies low-permeability segments. In this evolutionary system, two populations were maintained to evolve these two fuzzy rules simultaneously. The final fuzzy rule set were able to give sensible permeability interpretation for all well log segments. This work is reported in [7].

5 Conclusions

Well log interpretation is a routine, but time consuming, task in energy companies. With the increasing global energy demand, it is a natural trend to seek computerized well log interpretation techniques to provide results more efficiently. We have devised a method that maps the numerical well log data into a fuzzy symbol representation. This representation not only maintains the original well curve character but is more interpretable than its crisp numerical counter-part. The quality of this representation is verified using 2 sets of well logs data from offshore fields in South Africa. Based on the fuzzy symbolic representation, we have also implemented an intelligent well logs interpreter to interpret permeability with promising results [7]. We are currently applying this methodology to generate interpretation system for a wider ranges of reservoir properties, such as lithology and reservoir facies.

Acknowledgment

We thank Chevron for their permission to publish this work.

References

1. Abonyi, J., Feil, B., Nemeth, S., Arva, P.: Modified Gath-Geva Clustering for Fuzzy Segmentation of Multivariate Time-Series. Fuzzy Sets and Systems 149, 39–56 (2005)
2. Apostolico, A., Bock, M.E., Lonardi, S.: Monotony of Surprise and Large-Scale Quest for Unusual Words. In: Proceedings of the 6^{th} International Conference on Research in Computational Molecular Biology, pp. 22–31 (2002)
3. Asquith, G., Krygowski, D.: Basic Well Log Analysis, 2nd edn, American Association of Petroleum Geologists (2004)
4. Keogh, E., Chakrabarti, K., Mehrotra, S., And Pazzani, M.: Locally Adaptive Dimensionality Reduction for Indexing Large Time Series Databases. In: Proceedings of ACM SIGMOD Conference on Management of Data, pp. 151–162 (2001)
5. Larsen, R.J., Marx, M.L.: An Introduction to Mathematical Statistics and Its Applications, 2nd edn. Prentice Hall, Englewood Cliffs (1986)
6. Yu, T., Wilkinson, D., Xie, D.: A Hybrid GP-Fuzzy Approach for Reservoir Characterization. In: Riolo, R.L., Worzel, B. (eds.) Genetic Programming Theory and Practice, pp. 271–290. Kluwer, Dordrecht (2003)
7. Yu, T.: A Co-Evolutionary Fuzzy System for Reservoir Well Logs Interpretation. In: Yu, T., Davis, D., Baydar, C., Roy, R. (eds.) Evolutionary Computation in Practice. Springer, Heidelberg (2007)