Automating Human Reasoning with Natural Language Processing

R. Russell Rhinehart

Part I, CONTROL for the process industries, Vol. 34, No. 3, March, 2021, pp. 32-35
Part II, CONTROL for the process industries, Vol. 34, No. 4, April, 2021, pp. 37-39

Part 1

Introduction

Humans can drive a car: Keep it in the lanes, at the right speed, avoid obstacles, obey traffic signals, adjust heaters, switch wipers on and off, and monitor the display. Humans can cook on a grill, or run to a position in left field to catch a fly ball. Humans do not have mathematical models to perform such functions. Instead, they do the analysis and control action intuitively. In the control room, operators interpret and respond to many events; again, without mathematical models.

It would be nice to be able to automate many of those routine and intuitively executed monitoring and control activities. And we can, using several forms of artificial intelligence. This article will describe one approach originally labeled Fuzzy Logic (FL), but now often called Linguistic Rules, Natural Language Processing, or any of many alternate names that do not sound like faulty logic or uncertain reasoning. In 1965 FL was originally introduced by Lofti Zadeh (1921-2017), and to his mathematician and computer scientist community there was no undesirable connotations of the name. To make it palatable for those in the process industry, I'll use the term Natural Language Processing (NLP).

NLP Control (NLPC) is similar to using "IF-THEN" statements in automation. For example, "IF the discharge receiver is full, THEN switch to the other tank." However, such logical moves abruptly switch from one condition to another. The switch could be disruptive. Smooth transitions are desirable: In stopping a car we progressively increase breaking force then progressively lighten up when approaching the stop point. We do not instantly switch breaks to full on, then full off. NLPC permits the automation of smooth transitions.

NLPC also permits the automation of qualitative human reasoning. For instance, an automobile operator may look at the car speedometer and think the speed is a bit slow, and press on the accelerator pedal a bit harder. Or, if climbing a hill and pulling a trailer, the same speed error would be fixed by a medium-large pedal change. The human operator is an effective nonlinear controller, and NLPC can implement human qualitative assessment, such as "a bit slow", "a bit harder", or "a medium-large change". Humans give such imprecise instructions to each other, which are fully implementable by the listener. Qualitative assessments and direction are legitimate in control.

Natural Language Processing Control (NLPC) has demonstrated success in control related to chemical and mineral processes, robotics, electronic devices, optimization procedures, home appliances, credit approval, cement kiln control, anti-lock breaking, elevator scheduling, camera auto focus, and many other applications; and it offers many application benefits within the chemical process industry where human judgements can be routinized. Many of the control system vendors offer a NLPC product. The author has personally contributed to implementation of NLPC on commercial and lab-scale in-line pH control, correction of weekly plant mass flow measurements from plant-wide material balances, adjusting feed to eight parallel full-scale reactors based on temperature observation, adjusting lab-scale dynamic process model coefficients in real-time, and others.

This article presents the basic concepts, terminology, and the relatively simple mathematics of NLP and NLPC; hopefully, so that the reader will be able to determine when and how to use NLP for process analysis and NLPC for decision automation.

Crisp and Soft Categorization

The engineering and scientific mind-set has been programmed to value precise numerical values and crisp logic. However, the same scientists and engineers make personal life-important decisions, every minute, with qualitative perceptions and decision analysis. Examples include: When should I leave here to arrive there on time? How should I act when the employee comes in late again? Should I stop or go through the light that just turned yellow?

Although NLP is a more realistic representation of the human view of the world than crisp logic, we have been trained to think that Venn diagrams, and associated crisp logic, are the right way of describing the universe. The crisp logic concepts represent idealizations. They simplify analysis, but do not always represent the real world. To understand and utilize NLP, one must accept two concepts. One is the utility of qualitative assessment and decision, and the second is the pretense of crisp logic.

We act on qualitative understanding, and I'll use the following as an example in developing NLPC: One might observe that the outdoor temperature is very cold, but the wind is calm, and the sun is shining, and the activity plan for the day does not include much out-door time; and conclude, "Wear a light jacket". The linguistic terms "very cold", "calm", "shining", "much" are fully adequate descriptors of the variables to make an appropriate decision; and the set of variables "temperature", "wind speed", "sun intensity", and "time duration" are the complete set needed to make a decision – "light jacket".

NLP is a system for mathematical manipulation of such qualitative linguistic concepts. When one accepts that qualitative perceptions and assessments are a valid and sufficient basis for good decisions, one is accepting the utility of qualitative assessment.

Second, the pretense associated with precise quantification and calculations needs to be recognized. Yes, we want our chemical processes and automation decisions to be precise, accurate, exact, and scientifically grounded. Our enterprise success, our safety record, our employment, and our welfare depend on it. But what data is precise? Is it orifice-measured flow rate which could have a 2-5% noise amplitude related to fluid turbulence, and 5-10% bias related the calibration and the ideal square root assumption? Measurement devices are originally selected for "adequate accuracy", and then are calibrated to be "good enough". The FOPDT models that underly diverse control structures and tuning are just approximations to the reality.

When separate individuals tune a controller, for a perceived good enough performance using individually preferred tuning techniques, do they get the exact same values for gain, integral time, and derivative time?

The reality is that we let the control computers use uncertain numerical values, imprecise equation coefficient values, and approximate models to perform the automation decisions. In this view, they use wrong equations and noise-confounded data to take control action. Accepting this reality about precision removes a barrier to accepting NLP.

Some definitions are important: *Process variables* are those that describe the process situation or state. For process control these would be measurements (e.g. temperature, pressure, flow rate) or state or quality estimates from measurements (e.g. composition, molecular weight variance). In the example analysis of the outdoor weather to determine what to wear, these process variables were "temperature", "wind speed", "sun intensity", and "time duration". From the controller perspective, the process variables are the *input variables* for the decision procedure.

Conventionally, the term *linguistic category* is the term chosen to describe the process variables. Above, these were the adjectives and superlatives "very cold", "calm", "shining", "much".

Membership in Linguistic Categories

Venn diagrams are an idealization, which represents the universe of all objects as a rectangle. Within the universe rectangle is a circle that encloses the set of all objects in a particular category.

Consider an application that segregates apples from the universe of objects. Apples are inside the circle. Then take one bite from the apple. Is it still an apple? Or is it totally not an apple? Take more bites, until it becomes an apple core, which is not an apple. At what bite did it move from the is-apple category to the not-apple category? Ask someone who is eating an apple after the second bite, "What are you doing?" Will they answer, "I'm eating a not-apple"? The idealization of either being in the circle or out of the circle is fundamentally incorrect. At every bite there is a belongingness to the linguistic category of apple, and the belongingness fades from unity to zero with each bite.

NLP starts with a new definition of belongingness, illustrated in Figure 1 as the classification of outdoor temperature within the linguistic category of “hot”.

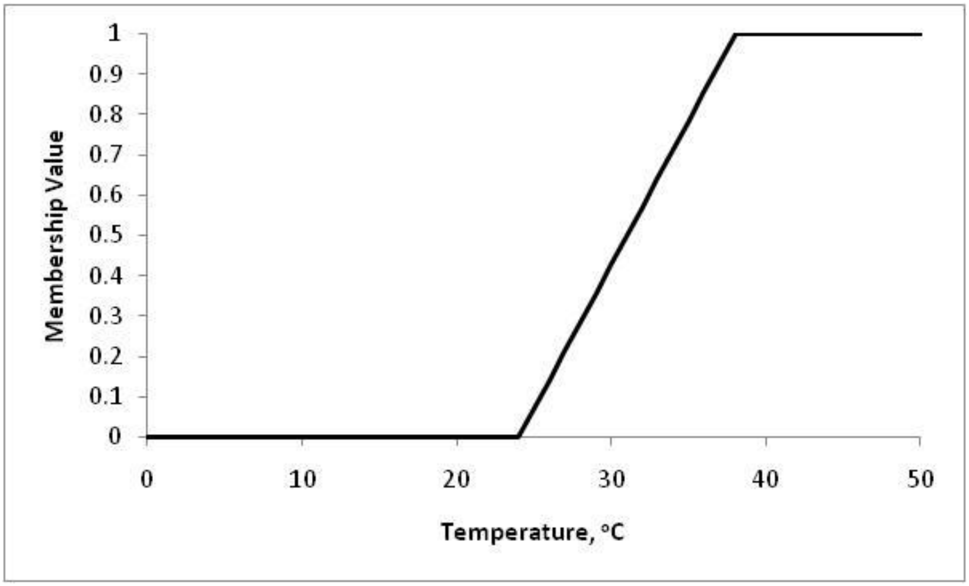

Figure 1 – Outdoor Temperature Membership Function in the Linguistic Category of Hot

The horizontal axis of Figure 1 presents a range of outdoor temperatures, and the vertical axis the belongingness of that temperature to the linguistic category “hot”, meaning an uncomfortable temperature. The curve represents my interpretation of whether a particular outdoor temperature is hot, or not. I think any value above 35°C (95°F) is uncomfortably hot, and any value below 24°C (75°F) is not uncomfortably hot. However, in between, it could be some of either, depending on my activity, or the sun, or the wind, or the humidity.

The variable represented on the horizontal axis in Figure 1 is outside temperature, this is termed the *process variable*. The *linguistic category*, hot, is the human description of the level of the variable. The *belongingness* of the process variable to the linguistic category, is represented on the vertical axis on a 0 to 1 basis. And the *membership function* is the “curve” on the figure that maps process variable to belongingness. This simple membership function “curve” is composed of three straight lines with *break points* at the process variable values of 24°C (75°F) and 35°C (95°F), which I defined as definitely in and definitely out of the linguistic category “hot”.

This curve seems to generally agree with my audiences. When I ask how many think 27°C (80°F) is hot, about 25% raise their hands. About 75% raise their hands when asked is 32°C (90°F) hot. But it seems very likely that those from hotter or colder climates would shift the curve to the left or right. The break points on the membership function reflect human opinion, which could be determined by any criterion that is relevant to the application. For instance, individuals who live nearer to the Equator might say that 30°C (85°F) is comfortable; while others who live near an Artic Circle might perceive the same temperature of 30°C (85°F) as uncomfortably hot.

Here are some approaches to determining break points:

- The fraction of people in agreement
- Opinion of one person
- Probability of acceptance, or un-detection
- The boss' opinion
- Fraction of capacity
- Fraction of sufficiency, function, perfection, or utility

The people who are handling the situation that you are considering for an NLP implementation have already thought about the right way to categorize the process variable into linguistic categories. They can state reasonable values for the linguistic categories and associated breakpoints that are appropriate to the application.

Membership functions have numerical values, as illustrated in Figure 1, which can easily be calculated from Equation Set (1), which represents the three line segments.

$$\left\{\begin{array}{lr} \mu_{HOT}(x) = 0, & x < 24^{o}C \\ \mu_{HOT}(x) = \frac{x-24}{35-24}, & 24^{o}C \leq x < 35^{o}C \\ \mu_{HOT}(x) = 1, & x \geq 35^{o}C \end{array}\right\} \quad (1)$$

Outdoor temperature can be labelled by other linguistic categories, and Figure 2 represents the use of three, Cold, Nice, and Hot. Equation Sets (2) & (3) show how to calculate the respective membership function values.

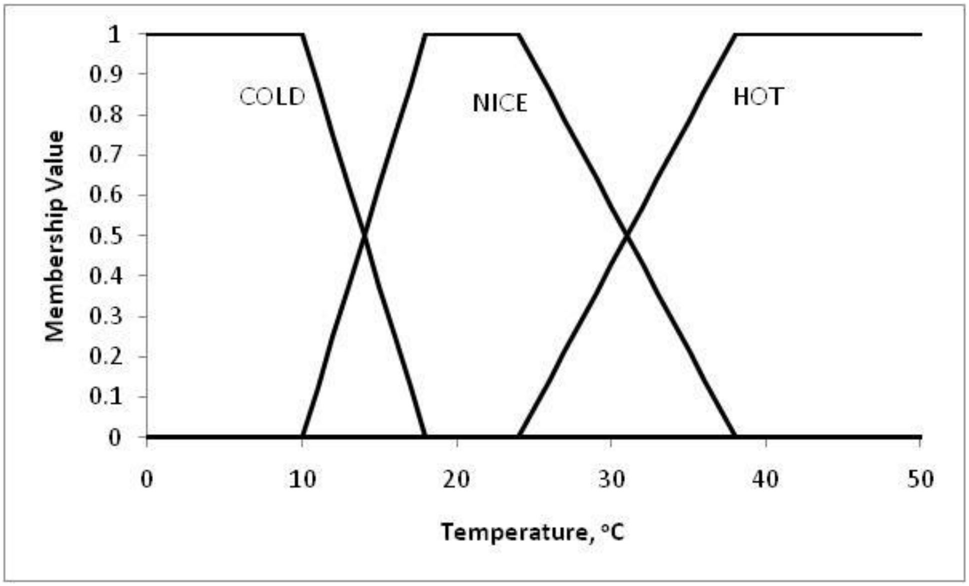

Figure 2 – Three Membership Functions for Outdoor Temperature

$$\begin{Bmatrix} \mu_{NICE}(x) = 0, & x < 10^oC \\ \mu_{NICE}(x) = \frac{x-10}{18-10}, & 10^oC \le x < 18^oC \\ \mu_{NICE}(x) = 1, & 18^oC \le x < 24^oC \\ \mu_{NICE}(x) = 1 - \frac{x-24}{35-24}, & 24^oC \le x < 35^oC \\ \mu_{NICE}(x) = 0, & x \ge 35^oC \end{Bmatrix} \quad (2)$$

$$\begin{Bmatrix} \mu_{COLD}(x) = 1, & x < 10^oC \\ \mu_{COLD}(x) = 1 - \frac{x-10}{18-10}, & 10^oC \le x < 18^oC \\ \mu_{COLD}(x) = 0, & x \ge 18^oC \end{Bmatrix} \quad (3)$$

There are several notable features in the figure. The membership function for the linguistic category "Nice" has four breakpoints, and a trapezoidal central shape. In this description, the membership function is not symmetric.

As is commonly practiced, each membership function has a range of 0 to 1 and a some overlap with the membership function of its adjacent linguistic categories. And the breakpoints are common to adjacent membership functions. For example, a temperature of 32°C (89°F) has about 0.4 *belongingness* to the linguistic category "Nice", 0.6 to "Hot", and 0.0 to "Cold".

Belongingness is a term used to interpret the meaning of the membership function value. A value of zero indicates that the process variable is definitely not in that linguistic category. A value of unity indicates that it is perfectly, exclusively, in that linguistic category. In-between values indicate the extent that it belongs to that linguistic category.

Because of the linear transition between breakpoints and the common break point locations, as illustrated in Figure 2, at any process variable value, the sum of all membership function values is unity. (There are alternate ways to represent membership functions. I like this simple approach.)

Outdoor temperature is one process variable that is used to determine what clothes to wear. This introduction to NLPC concepts will end with control action – choice of clothing. However, wind speed is another process variable that influences choice of clothing; so, it needs to be assessed. Figure 3 presents membership functions for three linguistic categories for wind speed, "Calm", "Breezy", and "Windy", and Equation Sets (4) through (6) show the mathematical relations. Notable features are that the membership functions do not have to be symmetric, and that the middle one can be triangular, not trapezoidal. The graph illustrates that a wind speed of 7 km/hr has a 0.1 belongingness to the linguistic category "windy" a 0.9 belongingness to "breezy", and a 0.0 belongingness to "calm". Again, because of the linear transition between breakpoints and the common break point locations, at any process variable value, the sum of all membership function values is unity.

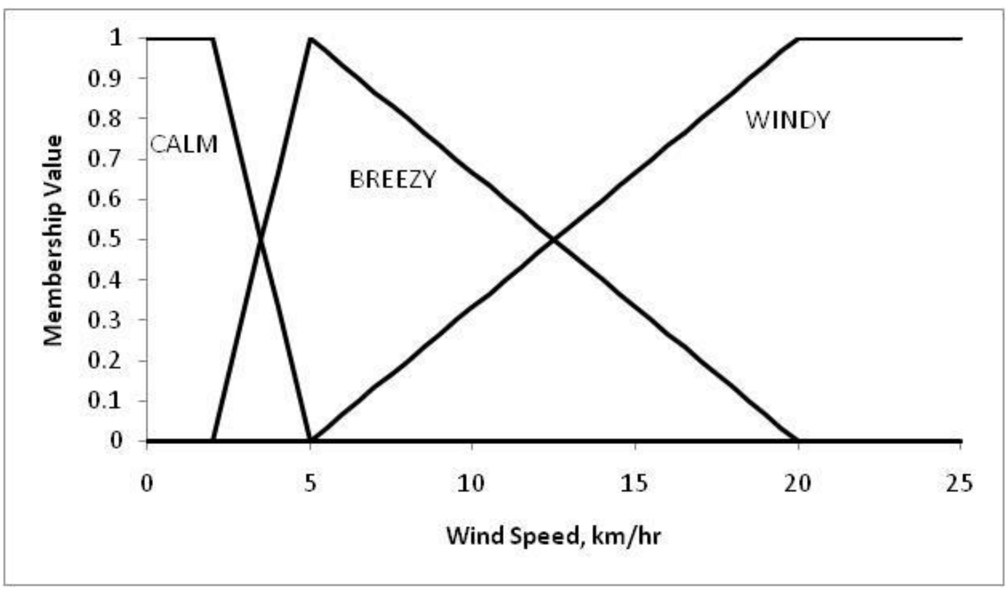

Figure 3 – Membership Function Values for Wind Speed

$$\begin{Bmatrix} \mu_{CALM}(x) = 1, & x < 2km/h \\ \mu_{CALM}(x) = 1 - \frac{x-2}{5-2}, & 2km/h \le x < 5km/h \\ \mu_{CALM}(x) = 0, & x \ge 5km/h \end{Bmatrix} \qquad (4)$$

$$\begin{Bmatrix} \mu_{NICE}(x) = 0, & x < 2km/h \\ \mu_{NICE}(x) = \frac{x-2}{5-2}, & 2km/h \leq x < 5km/h \\ \mu_{NICE}(x) = 1 - \frac{x-5}{20-5}, & 5km/h \leq x < 20km/h \\ \mu_{NICE}(x) = 0, & x \geq 20km/h \end{Bmatrix} \quad (5)$$

$$\begin{Bmatrix} \mu_{WINDY}(x) = 0, & x < 20km/h \\ \mu_{WINDY}(x) = \frac{x-5}{20-5}, & 5km/h \leq x < 20km/h \\ \mu_{WINDY}(x) = 1, & x \geq 20km/h \end{Bmatrix} \quad (6)$$

In these examples, each process variable only had three linguistic categories. One could extend the temperature descriptions to "Very Cold" and "Frigid", or to any number of intermediate categories. If there are many distinct linguistic categories in common use by the people that are analyzing and deciding about a process, this indicates that the categories are individually important, and should be included in an NLP description. However, only those descriptions that have distinctly different implications for the human analysis should be used.

Part 2 of this article will reveal how to use NLP for control and is scheduled for the MONTH (add the reference) issue of CONTROL magazine.

- - - - - - -

Russ Rhinehart started his career in the process industry. After 13 years and rising to engineering supervision, he transferred to a 31-year academic career. Now "retired", he enjoys coaching professionals through books, articles, short courses, and postings on his web site www.r3eda.com.

- -

Part 2

This article is Part 2 of a two-part article on Natural Language Processing. Part 1 appeared in (add reference).

Natural Language Processing Control (NLPC)

Control is the procedure of deciding and implementing action, it does not have to mean regulatory feedback algorithms. Here are some rules that we might use as control action, which are relevant to making a decision about what to wear:

If it is cold and breezy, then wear long pants, sweater, and jacket.
If it is nice and calm, then wear short-sleeved shirt and light-weight pants.
If it is cold and windy, then wear long pants, sweater, coat, hat, and gloves.

Very generally, the rule structure is:

IF (antecedent situation) THEN (control action)

And more specifically:

IF (Temperature is category AND Wind is category) THEN (Wear N items)

Since there are three linguistic categories for each of two process variables in the example for Figures 3 and 4 (from Part 1 of the article) there would be $3^2=9$ rules. Note that the rules are stated for the perfect or ultimate linguistic categories, but that most actual conditions will result in some validity to two adjacent temperature and two adjacent wind categories, making four of the nine rules simultaneously and partially valid.

Table 1 summarizes the nine rules.

Table 1 – Matrix of What to Wear

Wind Speed			
Windy	9 Items	6 Items	4 Items
Breezy	7 Items	5 Items	3 Items
Calm	6 Items	5 Items	3 Items
Temperature	Cold	Nice	Hot

Here, one rule is of the form:

IF (Outdoor Temperature is in Category "Cold", and Wind Speed is in Category "Breezy")
THEN (the appropriate action is: Wear 7 items of clothes.)

This sort of logic can be translated to NLPC for process control. For a heat exchanger this might be specifically:

IF (Temperature actuating error is positive-low, and Rate of Change is negative-large) THEN (incrementally adjust the controller output by +0.5%)

The IF part of such rules is termed the *antecedent*, and the THEN part is the *consequent*.

Returning the discussion to the weather and clothing decision from Part 1 of the article, using a temperature of 32^{o}C (89^{o}F) and a wind speed of 7 km/hr. At these conditions, Figure 2 and associated equations show that the temperature membership values are 0.0, 0.6, and 0.4. This is an in-between classification from “Nice” to “Hot”, mostly “Nice”. The temperature has a 0.6 belongingness to the “Nice” category, which indicates a 0.6 belongingness to the central column in Table 1 (and 2). Temperature has a 0.4 belongingness to the category “Hot” and to the right-hand column, and a 0.0 belongingness to the category “Cold” and to the left-hand column. Similarly, Figure 3 and associated equations show that the wind speed of 7 km/hr has a 0.0 belongingness to the Calm category and the lower row of Table 2. It also has a 0.9 belongingness to the Breezy category and the middle row, and a 0.1 belongingness to the Windy category and its upper row.

There are two common methods to choose the belongingness to each cell in the Table 2 matrix. For the first, consider belongingness as a probability or likelihood, then the probability that the condition is in one particular cell is the probability that it is in the row of the cell AND in the column of the cell. (Although, useful, the likelihood analogy is weak.) But with the likelihood viewpoint, products of row-column belongingness values represent the *truth* that a process variable set is in a particular cell. These cell values are also indicated in Table 2.

Table 2 - Matrix of Dress Rules Showing Rule Truth

Windy μ_W(7 km/h)=0.1	9 Items Truth = 0.1*0 = 0	6 Items Truth = 0.1*0.6 = 0.06	4 Items Truth = 0.1*0.4 = 0.04
Breezy μ_B(7 km/h)=0.9	7 Items Truth = 0.9*0 = 0	5 Items Truth = 0.9*0.6 = 0.54	3 Items Truth = 0.9*0.4 = 0.36
Calm μ_C(7 km/h)=0	6 Items Truth = 0*0 = 0	5 Items Truth = 0*0.6 = 0	3 Items Truth = 0*0.4 = 0
	Cold μ_C(32°C)=0	Nice μ_N(32°C)=0.6	Hot μ_W(32°C)=0.4

Conveniently, the sum of all of the truth values is unity because 1) the membership functions are linear, 2) adjacent membership functions share break points, and 3) the row-column membership product is used for the rule truth. (There are alternate approaches. I prefer this simple one.)

The truth of a rule is the weight or importance given to that rule, and the blended control action is the truth-weighted sum of all rules, as described in Equation (7) and illustrated for this example in Equation (8). Here T_k represents the truth of the k^{th} rule, and A_k represents the action to be taken if the k^{th} rule was perfectly true.

$$Action = \sum_{AllRules} T_k A_k \qquad (7)$$

$$Action = 0 \cdot 9 + 0.06 \cdot 6 + 0.04 \cdot 4 + 0 \cdot 7 + 0.54 \cdot 5 + \\ +0.36 \cdot 3 + 0 \cdot 6 + 0 \cdot 5 + 0 \cdot 3 = 4.3 \ \ Items \tag{8}$$

The calculated action "Wear 4.3 items" is the control action. If this was a controller output command to a valve, "Open 4.3 % more", then the decimal part is acceptable. However, because the answer in this clothing example can only be an integer, one might round the result as is done in digital processing with a finite bit length for variables.

Equation (7), and its example in Equation (8), converts the qualitative rules and qualitative characterization of the process variables into a definitive implementable value.

Another common convention to assess the truth of a rule is to use the minimum of the row-column membership values as the truth for the cell. In general then, the truths do not sum to unity, so an additional weighting uses the individual truths normalized by the sum of all truths. In my experience the several approaches are equivalent.

The extension to analysis and control of higher dimension situations is straight forward. For example, relative humidity could be included as a third consideration in making the choice of what to wear. This would place three process variables in the antecedent making it of dimension three. For a 3-D antecedent, the rule matrix would be a rectangular volume. If sun intensity and time duration are also considered in the decision process the antecedent would have 5 dimensions. The extension to higher-order antecedents is easily performed in programming, but it is not amenable to visualization.

Comments

NLPC implements human rules for automating decisions, without advanced mathematics (such as calculus or Laplace transforms), and it permits nonlinear action. If you have applications in which IF-THEN conditionals automate the routine engineering and operator decisions or actions, then you are almost implementing NLPC. Better than the normal use of IF-THEN conditionals, NLPC permits smooth transitions between categories.

NLP provides a framework for standardization of how heuristic rules are implemented. Operators and managers can understand the linguistic logic, and they can state the rules in natural language. The recorded body of rules provides additional benefits in training and developing process understanding.

NLP is simple to implement and document.

Normally, process experience among operators and engineers is sufficient to develop an NLPC application without performing either experimental process response testing or controller tuning explorations. NLPC can integrate user-defined action for feedforward and constraint avoidance.

However, NLPC needs break points defined, perhaps averaging three for the first linguistic category, then one more for each additional category for each process variable, and one control action for each rule. If NLPC is replacing PI control, then commonly there are two process variables, actuating error and rate of change of error. If each has 5 linguistic categories (zero and two "+" and two "-" categories), then there are 25 rules and about 14 break points, summing to about 40 user-required values. Gain scheduled PI control over 4 regions needs 8 tuning values and 3 break points for a total of 11 values. The simplicity of gain-scheduled PI control, and the widespread familiarity with PI algorithms, might override the benefits of using NLPC to replace PI for feedback control. Most NLPC vendors offer PI substitutes and software features that make it simple to setup improved control. But in my opinion, replacing PI in feedback control is <u>not</u> where NLPC has its largest advantages.

NLPC should be considered as an automaton solution wherever engineers or operators observe, perceive, and take routine supervisory corrective action. Consider NLPC for automating process management action, rather than replacing feedback control. If you are either personally implementing or automating the implementation of heuristic rules, you have a potential application for NLPC. Identify where your people are routinely observing something and taking corrective action, then consider automating that action with NLPC. For example: Do they observe time-to-breakthrough on a carbon-bed absorber to adjust the absorb-to-steaming cycle period? Do they observe outlet composition on parallel reactors to adjust feed rate between reactors? Do they observe zero-crossing behavior of a controller to increase or decrease gain? Using the NLPC structure will standardize the multiple applications of heuristic supervisory activities throughout your enterprise.

As with any automation approach, the control rules and the categorization of process variables in NLPC reflect the knowledge of the creator, not necessarily the best logic, and they may even integrate folklore. That it works, does not mean it represents either a valid or best underlying understanding. Use an application to see if results affirm your intuitive understanding. Be willing to improve.

Further, like any control strategy or controller tuning, NLPC reflects the process understanding at one time, which may need to be revised when the process equipment is changed or used under significantly different conditions.

Finally, don't call it Fuzzy Logic! Dr. Zadeh's insight on logic and its mathematical formulation was transformational, simple, and effective. He gets much respect! But his choice of that term does not evoke confidence or security in the minds of process managers. If you want to implement it within industry, call it Natural Language Processing, or whatever makes it acceptable in your community.

- - - - - - -

Russ Rhinehart started his career in the process industry. After 13 years and rising to engineering supervision, he transferred to a 31-year academic career. Now "retired", he enjoys coaching professionals through books, articles, short courses, and postings on his web site www.r3eda.com.