Fuzzy Systems Are Universal Approximators for Random Dependencies: A Simplified Proof

Mahdokht Afravi and Vladik Kreinovich

Department of Computer Science, University of Texas at El Paso
El Paso, TX 79968, USA
mafravi@miners.utep.edu, vladik@utep.edu

Abstract. In many real-life situations, we do not know the actual dependence $y = f(x_1, \ldots, x_n)$ between the physical quantities x_i and y, we only know expert rules describing this dependence. These rules are often described by using imprecise ("fuzzy") words from natural language. Fuzzy techniques have been invented with the purpose to translate these rules into a precise dependence $y = \widetilde{f}(x_1, \ldots, x_n)$. For deterministic dependencies $y = f(x_1, \ldots, x_n)$, there are universal approximation results according to which for each continuous function on a bounded domain and for every $\varepsilon > 0$, there exist fuzzy rules for which the resulting approximate dependence $\widetilde{f}(x_1, \ldots, x_n)$ is ε-close to the original function $f(x_1, \ldots, x_n)$.
In practice, many dependencies are *random*, in the sense that for each combination of the values $x_1, \ldots, x_n$, we may get different values y with different probabilities. It has been proven that fuzzy systems are universal approximators for such random dependencies as well. However, the existing proofs are very complicated and not intuitive. In this paper, we provide a simplified proof of this universal approximation property.

1 Formulation of the Problem

It is important to determine dependencies. One of the main objectives of science is to find the state of the world and to predict the future state of the world – both in situations when we do not interfere and when we perform a certain action. The state of the world is usually characterized by the values of appropriate physical quantities.

For example:

– we would like to know the distance y to a distant star,
– we would like to predict tomorrow's temperature y at a given location, etc.

In some cases, we can directly measure the current value of the quantity y of interest. However, in many practical cases, such a direct measurement is not possible – e.g.:

– while it is possible to measure a distance to a nearby town by just driving there,

– it is not yet possible to directly travel to a faraway star.

And it is definitely not possible to measure tomorrow's temperature y today.

In such situations, since we cannot directly measure the value of the desired quantity y, a natural idea is:

– to measure related easier-to-measure quantities $x_1, \ldots, x_n$, and then
– to use the known dependence $y = f(x_1, \ldots, x_n)$ between these quantities to estimate y.

For example, to predict tomorrow's temperature at a given location, we can:

– measure today's values of temperature, wind velocity, humidity, etc. in nearby locations, and then
– use the known equations of atmospheric physics to predict tomorrow's temperature y.

In some cases we know the exact form of the dependence $y = f(x_1, \ldots, x_n)$, but in many other practical situations, we do not have this information. Instead, we have to rely on experts who often formulate their rules in terms of imprecise ("fuzzy") words from natural language.

Imprecise ("fuzzy") rules and how they can be transformed into formulas. What kind of imprecise rules can we have? In some cases, the experts formulating the rule are imprecise both about x_i and about y. In such situations, we may have rules like this: "if today's temperature is very low and the Northern wind is strong, the temperature will remain very low tomorrow." In this case, x_1 is temperature today, x_2 is the speed of the Northern wind, y is tomorrow's temperature, and the properties "very low" and "strong" are imprecise.

In general, we have rules of the type

"if x_1 is A_{k1}, ..., and x_n is A_{kn}, then y is A_k",

where A_{ki} and A_k are imprecise properties.

It is worth mentioning that in some cases, the information about x_i is imprecise, but the conclusion about y is described by a precise expression. For example, in non-linear mechanics, we can say that when the stress x_1 is small, the strain y is determined by a linear formula $y = k \cdot x_1$, with known k, but when the stress is high, we need to use a nonlinear expression $y = k \cdot x_1 - a \cdot x_2^2$ with known k and a. Here, both expressions are exactly known, but the condition when to apply one or another is described in terms of imprecise words like "small".

To transform such expert rules into a precise expression, Zadeh invented fuzzy logic; see, e.g., [1, 4, 5]. In fuzzy logic, to describe each imprecise property P, we ask the expert to assign, to each possible value x of the corresponding quantity, a degree $\mu_P(x)$ to which the value x satisfies this property – e.g., to what extent the value x is small. We can do this, e.g., by asking the expert to mark, on a scale from 0 to 10 to what extent the given value x is small. If the expert marks 7, we take $\mu_P(x) = 7/10$. The function $\mu_P(x)$ that assigns this degree is known as the *membership function* corresponding to the property P.

For given inputs $x_1, \ldots, x_n$, a value y is possible if it fits within one of the rules, i.e., if:

– either the first rule is satisfied, i.e., x_1 is A_{11}, …, x_n is A_{1n}, and y is A_1,
– or the second rule is satisfied, i.e., x_1 is A_{21}, …, x_n is A_{2n}, and y is A_2, etc.

Since we assumed that we know the membership functions $\mu_{ki}(x_i)$ and $\mu_k(y)$ corresponding to the properties A_{ki} and A_k, we can thus find the degrees $\mu_{ki}(x_i)$ and $\mu_k(y)$ to which each corresponding property is satisfied.

To estimate the degree to which y is possible, we must be able to deal with propositional connectives "or" and "and", i.e., to come up with a way to estimate our degrees of confidence in statements $A \vee B$ and $A \,\&\, B$ based on the known degrees of confidence a and b of the elementary statements A and B. In fuzzy logic, such estimation algorithms are known as *t-conorms* ("or"-operations) and *t-norms* ("and"-operations). We will denote them by $f_\vee(a, b)$ and $f_\&(a, b)$. In these terms, the degree $\mu(y)$ to which each value y is possible can be estimated as $\mu(y) = f_\vee(r_1, r_2, \ldots)$, where

$$r_k \stackrel{\text{def}}{=} f_\&(\mu_{k1}(x_1), \ldots, \mu_{kn}(x_n), \mu_k(y)).$$

We can then transform these degrees into a numerical estimate $\overline{y}$. This can be done, e.g., by minimizing the weighted mean square difference $\int \mu(y) \cdot (y - \overline{y})^2\, dy$, which results in

$$\overline{y} = \frac{\int y \cdot \mu(y)\, dy}{\int \mu(y)\, dy}.$$

Universal approximation result for deterministic dependencies. For deterministic dependencies $y = f(x_1, \ldots, x_n)$, there are universal approximation results according to which for each continuous function on a bounded domain and for every $\varepsilon > 0$, there exist fuzzy rules for which the resulting approximate dependence $\widetilde{f}(x_1, \ldots, x_n)$ is ε-close to the original function $f(x_1, \ldots, x_n)$ for all the values x_i from the given domain.

In practice, we can often only make probabilistic predictions. In practice, many dependencies are *random*, in the sense that for each combination of the values $x_1, \ldots, x_n$, we may get different values y with different probabilities.

Fuzzy systems are universal approximators for random dependencies as well. It has been proven that fuzzy systems and universal approximators for random dependencies as well; see, e.g., [2, 3].

Remaining problem: can we simplify these proofs. The proofs presented in [2, 3] are very complicated and not intuitive. It is therefore desirable to simplify these proofs.

What we do in this paper. In this paper, we provide a simplified proof of the universal approximation property for random dependencies.

2 Towards a Simplified Proof

Main idea: how do we simulate random dependencies? To simulate a deterministic dependence $y = f(x_1, \ldots, x_n)$, we design an algorithm that, given the values $x_1, \ldots, x_n$, computes the corresponding value y.

To simulate a random dependence, a computer must also use the results of some *random number generators* that generate numbers distributed according to some probability distribution. Such generators are usually based on the basic random number generator – which is either supported by the corresponding programming language or even on a hardware level – that generates numbers uniformly distributed on the interval $[0, 1]$.

From this viewpoint, the result of simulating a random dependency has the form

$$y = F(x_1, \ldots, x_n, \omega_1, \ldots, \omega_m),$$

where F is the corresponding algorithm, x_i are inputs, and the values ω_j comes from the basic random number generator.

In these terns, what does it mean to approximate? In the above terms, to approximate means to find a function $\widetilde{F}$ for which, for all possible inputs x_i from the given bonded range, and for all possible values ω_j, the corresponding value

$$\widetilde{y} = \widetilde{F}(x_1, \ldots, x_n, \omega_1, \ldots, \omega_m)$$

are ε-close to the results of applying the algorithm F to the same values x_i and ω_j.

This leads to a simplified proof. The above idea leads to following simplified proof:

- due to the universal approximation theorem for deterministic dependencies, for every $\varepsilon > 0$, there exists a system of fuzzy rules for which the value of the corresponding function $\widetilde{F}$ is ε-close to the value of the original function F;
- thus, we get a fuzzy system of rules that provides the desired approximation to the original random dependency F.

Acknowledgments

This work was supported in part by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721, and by an award "UTEP and Prudential Actuarial Science Academy and Pipeline Initiative" from Prudential Foundation.

References

1. G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.

2. P. Liu, “Mamdani Fuzzy System: Universal Approximator to A Class of Random Processes”, *IEEE Transactions on Fuzzy Systems*, 2002, Vol. 10, No. 6, pp. 756–766.
3. P. Liu and H. Li, “Approximation of stochastic processes by T-S fuzzy systems”, *Fuzzy Sets and Systems*, 2005, Vol. 155, pp. 215–235.
4. H. T. Nguyen and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2006.
5. L. A. Zadeh, “Fuzzy sets”, *Information and Control*, 1965, Vol. 8, pp. 338–353.