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

1. Main Objective of Science in General

- One of the main objectives of science is:
  - to find the state of the world and
  - to predict the future state of the world.
- We need to do it both:
  - in situations when we do not interfere and
  - in situations 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.

2. Direct Measurement Is Not Always Possible

- 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.

3. Need for Indirect Measurements

- 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.

4. It Is important to Determine Dependencies

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

5. Imprecise ("Fuzzy") Rules

- 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.

6. Fuzzy Rules: General Case

- In general, we have rules of the following type, where A_{ki} and A_k are imprecise properties:

$$\text{“if } x_1 \text{ is } A_{k1}, \ldots, \text{ and } x_n \text{ is } A_{kn}, \text{ then } y \text{ is } A_k\text{”},$$

- 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 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.

7. Fuzzy Logic

- To transform such expert rules into a precise expression, Zadeh invented fuzzy logic.
- 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 0-to-10 scale, to what extent the given 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* corr. to P.

8. Fuzzy Logic (cont-d)

- 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.
- We assumed that we know the membership functions $\mu_{ki}(x_i)$ and $\mu_k(y)$ corresponding to 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 "or" and "and".

9. Need for "And"- and "Or"-Operations

- In other words, we need 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.
- 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 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 $\overline{y}$.

10. Fuzzy Technique: Final Step

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

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

11. Universal Approximation Result for Deterministic Dependencies

For deterministic dependencies $y = f(x_1, \ldots, x_n)$, there is the following universal approximation result:

- for each continuous function $f(x_1, \ldots, x_n)$ on a bounded domain D, 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 $f(x_1, \ldots, x_n)$ for all $(x_1, \ldots, x_n) \in D$:

$$|\widetilde{f}(x_1, \ldots, x_n) - f(x_1, \ldots, x_n)| \le \varepsilon.$$

12. Often, We Can Only Make Probabilistic Predictions

- In practice, many dependencies are *random*:
  - 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 random dependencies as well.
- The existing proofs are very complicated and not intuitive.
- It is therefore desirable to simplify these proofs.
- We provide a simplified proof of the universal approximation property for random dependencies.

13. Main Idea: What Is Random Dependence from an Algorithmic Viewpoint

- 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 $y = f(x_1, \ldots, x_n)$.
- To simulate a random dependence, we must also use the results of *random number generators* (RNG).
- Such a generator is usually based on the basic RNG that generates numbers ω_j uniform on $[0, 1]$.
- From this viewpoint, the result of simulating a random dependency has the form

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

14. In These Terms, What Does It Mean to Approximate?

- We have a random dependence

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

- To approximate means to find a function $\widetilde{F}(x_1, \ldots, x_n, \omega_1, \ldots, \omega_m)$ for which:
  - for all possible inputs x_i from the given bonded range, and
  - for all possible values ω_j
  - the result of applying $\widetilde{F}$ is ε-close to the desired value y:

$$|\widetilde{F}(x_1, \ldots, x_n, \omega_1, \ldots, \omega_m) - F(x_1, \ldots, x_n, \omega_1, \ldots, \omega_m)| \le \varepsilon.$$

15. This Leads to a Simplified Proof

- Let us apply the universal approximation theorem for deterministic dependencies.
- It implies that:
  - 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.
- The universal approximation result for random dependencies is thus proven.

16. 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 from Prudential Foundation.

17. Bibliography: Current Proofs of Universal Approximation Result for Random Processes

- 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.
- P. Liu and H. Li, "Approximation of stochastic processes by T-S fuzzy systems", *Fuzzy Sets and Systems*, 2005, Vol. 155, pp. 215–235.

18. General Bibliography on Fuzzy

- G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.
- H. T. Nguyen and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2006.
- L. A. Zadeh, "Fuzzy sets", *Information and Control*, 1965, Vol. 8, pp. 338–353.